ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/359056287

Health care counselling via voicebot using multinomial naive bayes algorithm

Conference Paper · March 2022

CITATION
1

READS
139

1 author:



Revathy Ramesh
Sathyabama Institute of Science and Technology
7 PUBLICATIONS **33** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

gestational diabetes research View project

Machine Learning in Healthcare Domain View project

All content following this page was uploaded by Revathy Ramesh on 07 March 2022.

The user has requested enhancement of the downloaded file.

Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020)
IEEE Conference Record # 48766; IEEE Xplore ISBN: 978-1-7281-5371-1

Health Care Counselling Via Voicebot Using Multinomial Naive Bayes Algorithm

S.Revathy
Associate Professor
School of Computing
Sathyabama Institute of Science And Technology
Chennai, Tamil Nadu, India
ramesh.revathy@gmail.com

Niranjani.R
Department of Information Technology
Sathyabama Institute of Science And Technology
Chennai, Tamil Nadu, India
nirusanjay24@gmail.com

Roslin Kanushya.J
Department of Information Technology
Sathyabama Institute of Science And Technology
Chennai, Tamil Nadu, India
roslinkanushya@gmail.com

ABSTRACT- Voicebot bestow the user, the ability of receiving retort within minutes, no matter their geographic location, or the apparatus, they are using to transmission. Data related to the health issues are gathered from medical experts, and all of them are stored on the server side of the system. The software furnished by the system offers answers and solutions for audit and treating patients after combining descriptions of the illness and the symptoms, diagnostics, predictions and prescription. This software will be greatly useful to people to protect them from complication, in this proposed system the Voicebot is used to solve the difficulties facing by people at the time of illness, which influence the health of the person. With all these factors, the Voicebot updates its knowledge base and allows patients to benefit from using it. This designed bot is commonly used for general Health Care. Here the input is given by the user in format of text. This text will be analyzed by Multinomial Algorithm, by which the data are fetched from the Data Base and generates voice response. When examine all these bots, the oddity of the proposed system can get response by both sensory and textual methods. This graces the untaught customers to use proficiently. It bears the whole system in customer's cordial manner. The need of the proposed system is minimum time and energy and can clarify our doubts at any time.

Keywords: Multinomial algorithm, NLP, Healthcare, VoiceBot

I. INTRODUCTION

A voicebot is a hunk of software that conducts a conversation via sensory or textual methods. A voicebot uses artificial intelligence (AI) software that can implicate a conversation with a user in natural language through messaging applications. Voicebot's provide the user, the ability of receiving replies within minutes, no matter their geographic location, or the device they are using to communicate. Improves customer service. Personalize communication. Automated repetitive tasks.

This software is created for general medicine like sugar, preeclampsia, blood pressure, fever and cold. Voicebot respond to the quick crisis. Uneducated people can also benefit with this software. It is the ongoing health care monitoring device. It provides support and additional information. People find it difficult to consult doctor each and every time during the time of deficiency. This software will be useful for them to clarify the queries related to some common disease like blood pressure and general medicine. With this software can clarify our doubts at any time, hence it need minimum time and energy. In this proposed system the Voicebot is used to clear and clarify the doubts and queries of patients during emergency. People may find it difficult to go to hospital each and every time when they are in need. Voicebot's provide the user, the ability of receiving retort within minutes, no matter their geographic location, or the apparatus they are using to transmission. The software provided by the system offers answers and solutions for audit and treating patients after combining descriptions of the illness and the symptoms, diagnostics, predictions and prescription. The impartial of the proposed system is quickly crisis, uneducated people also gratuity with this system, Scheduling rendezvous.

Ada is one of the most leading Health Care Voicebot. If symptoms to the bot is given, it will give you a possible diagnosis. It also helps you to book an appointment. Preeclampsia Voicebot, another leading Voicebot is specially designed for pregnant women who are all having preeclampsia. When comparing to all these bots, the uniqueness of the proposed system is that can interact with the bot via speech. This designed bot is commonly used for general Health Care.

The objectives are quick crisis, assist the untaught people also benefit with this software, assist the availability and ongoing health care monitoring, scheduling appointment and providing support and additional information.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 06,2022 at 10:17:42 UTC from IEEE Xplore. Restrictions apply.

II. RELATED WORKS

Luliana Marin et al [1] has developed Chabot for women to understand preeclampsia and its complications. It is the first Chabot for preeclampsia and its complications. In this system data about preeclampsia is collected from clinical experts, as well as from audited patients and all of them are stored on the server side of the system. All the triples are situated inside the RDF-ontology. The triples expressed the clinically validated knowledge which is verified by experts in order to find out if the notions have been well linked. The chatbot related queries are stored in SPARQl database which is not reliable.

A.Folstad, et al [2] has designed Chatbot through which the human and computer can interact .It creates revolution in Chabot. It is like a basic interaction between people. Producing a new quantum (knowledge base KB related data) that avoids overlapping, finding the personal questions and rejects the inadmissible queries by combining available QA and sensory formats.

F.T.Spradley [3] has developed chatbot which act as an advisor for hypertensive patients. It describe briefly about preeclampsia using ViDi (virtual validation of given queries using generic algorithms) technology. This system has the efficacy to remember the previous conversation and in order to work as a virtual adviser for hypertensive patients.

Braun, A.H.Mendez [4] has developed Estimating Natural Language understanding service for Conversational Question Answering System. It describes the conversational system and how to handle the queries. Redesigning and extending the sensory system by merging the expedient matching techniques in order to accelerate the sensory manner rather than QA form.

G.Leifer, et al [5] has developed a chatbot that take care of women with complications during pregnancy. It describes the nursing care and the complications faced by the pregnant women. Creating a new identical techniques OMAMC (One-match-All-match-category) in order to produce enhance results by reducing identical time and improving sensory flexibility.

Eucoor.S.J.du.Preez [6] has designed an intelligent web-based voice chatbot. It is a first chat bot which will display the queries in order. Intending a new system by using AIML matching pattern category. This system has SQL relational database to store the queries related to IES, In order to improve the transition between the consumers and dealers.

K.Bala sahila Pandita [7] has developed chatbot for management using AI, Sentimental analysis were done in this product. This system is used to enable the disabled people via voice. Intended the system using Edger Matching technique this is the combination of generic and tfd algorithms, doesn't contain any database to store data.

Agnese et al Augello, [8] has designed An Approach to enhanced chatbot semantic, maintainability, experiences with the FRASI projects. This algorithm combines the query identical and indexing methods with pattern matching and reflects to Information Retrieval techniques to create a new sentence pattern.

Emanule Haller, [9] has developed a chatbot that simulates a historical figure, Analyses the querying technology. There are two main categories are a, Transmute the two sentence into one b, Applying IR matching techniques to chatbot, developed by generic and edger matching algorithm.

Haller tac [10] has designed chatbot for imperative, Personalization and Utilitarian Design. This system produces new sentence pattern from two existing pattern. The study proposes exploiting an OMAMC algorithm to build a new sentences contingent on the sentence that are retrieved from available database.

Eucoor.S.J.du.Preez [11] has designed the chatbot of Ada, which is one of the most leading Health Care Chatbot. If give symptoms to the bot, it will give you a possible diagnosis. It also helps you to book an appointment. Preeclampsia chatbot, another leading chatbot is specially designed for pregnant women who are all having preeclampsia. When comparing to all these bots, the uniqueness of the proposed system is that can interact with the bot via speech.

R. Mustafa, et al [12] has developed chatbot which act as a hypertensive patients. It describes briefly about the hypersion infected disease technology. This system helps to identify the person who are suffering from hypertension.

E.Bartsch, et al [13] has designed a chatbot that take care of women with complications during pregnancy. It describes the nursing care and the complications faced by the pregnant women. Creating a new identical techniques of AIML.

I. Marin et al [14] has developed the chatbot for sensory system by merging the expedient matching

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 06,2022 at 10:17:42 UTC from IEEE Xplore. Restrictions apply.

techniques in order to accelerate the sensory manner rather than QA
Emanule Haller et al [15] has designed a chatbot for identify the figures or pattern matching. . Creating a new identical techniques OMAMC (One-match-All-match-category) in order to produce enhance results by reducing identical time and improving sensory flexibility.

S.Revathy et al [16] has validated rough fuzzy cluster algorithm using the technique known as Naïve-Bayesian.

III. PROPOSED SYSTEM

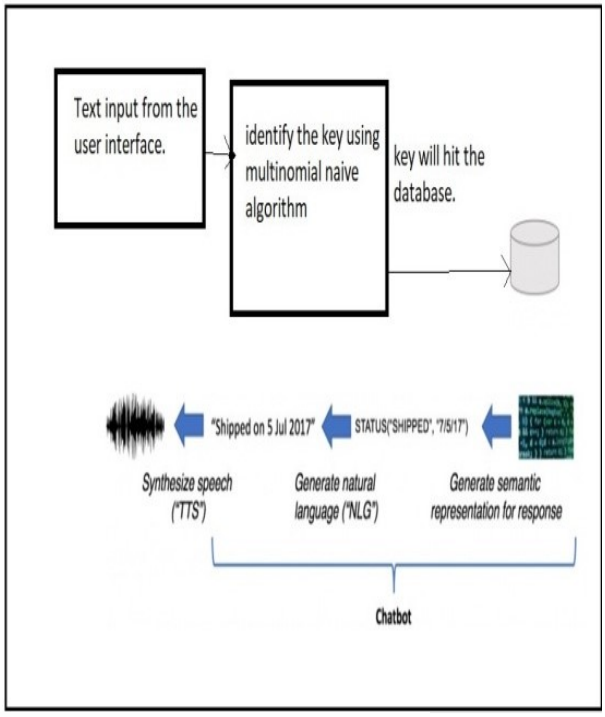
Methodology for Voicebot

A Voicebot using Multinomial Naïve Bayes Algorithm is desinged. It is used to analyze the queries given by the user and find the probability, which acts as a key that hits the Database to get an output.

Multinomial-Naive-Bayes

Naive Bayes classifier is kindred of probalistic algorithms based on employing Bayes theorem with the naïve premise of conditional independence betwixt every pair of constituent**.** Bayes theorem enumerates probability P (c|x).

$$P(c|x) = P(x|c) * P(c) / P(x) \quad (1)$$



.

Fig 1: Architecture of the Voicebot

Where c is the caliber of the possible outcomes, X is the given instance which has to be categorized, depicting some certain constituent

P (Positive): It helps to identify the probability of positive words in the tag of sentences which have been entered by the users.

P (Negative): It helps to identify the probability of negative words in the tag sentences which have been entered by the users**.**

In this process, the given input will be analysed by the listener trainer function. It will analyse the complete text and split that text by tokenization and stemming process which is the algorithm that assembles all the harmonized forms of a word in order to rend them down to their root dictionary form or lemma. Words are neglected into a part of speech by way of the rules of grammar. With reference to the database it will decide the key to retrieve the data. For example if the user gives the query for cold, the predefined libraries set the key as cold**.** Then the system will split the given input into separate unit and stores the data in dictionary format .

Applying multinomial algorithm in proposed system
Step1: Calculate the prior probabilities these are
The Probability of data being in a specific category from the given set of data.

P (positive category input) = (Number of positive data classified into a key) / (total data of Sentence pattern).
P (Negative category input) = (Number of Negative Data classified into a key) / (total Data of sentences pattern).

P (number of positive) =Number/total
P (number of negative) =Number/total
Step2: Calculate Likelihood by using conditional probability.

P (Word/Categorical data) = (Number of occurrence of the word in all tables from separated data) / (All the words + total number of unique words in all the data in database)
P (Category/data) =P (categorical data) X P (word1/data) X P (word2/data) X P (word N /data)
This algorithm methodology is effective for documentation problem.
The equation (1) is used to find the probability of words.
After identifying the probability, it will hit the database for the output. Before that the database should be

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 06,2022 at 10:17:42 UTC from IEEE Xplore. Restrictions apply.

stored in yml (Ain't Markup Language) format, this helps to make the database to human readable format. The output which is stored in form of text will be converted into voice format with the help of some predefined function. That function will analyze the output first, then it will convert each words into voice format by stemming process .In the algorithm there is function called engine that will check whether the given query is suitable for the user's text or not. After identifying it is a suitable query it will wait for sometimes for argument analysis. At last the output will displayed in both text and voice format. This algorithm is intended for text documents. It is a explicitly models the word counts and alters the underlying calculations to covenant with in.It evaluates the conditional probability of a particular word give a class as the relative frequency of term t in documents belonging to class(c).The variation takes into account the number of occurrences of term t in training documents from class(c), including multiple occurrences. The table 1 describes the training set of sample queries and table 2 shows the probability finding.

Table 1: Sample Training data set

Contents	Category
Suddenly I got faint	Low sugar
What can I do for low sugar	Low sugar
Having high sugar	High sugar

Table 2: Conditional probability for keys.

Word	occurrences	Total words	Conditional probability	Total uniqueness
Faint	1	4	0.00234	2
Low sugar	1	6	0.1666	2
High sugar	1	3	0.1052675	2
(stress	1	5	0.052631	2

Likewise it will find the probability for each word in the sentences .Before finding the probability, it will identify the positive and negative occurrences of the words which have been entered by the users using stemming and tokenization process.

PROPOSED ALGORITHM

Step 1: Download the Voicebot related libraries and database which is stored in yml (Ain't Markup Language Format).This format makes the data as human readable format.

Step 2: Once the input is given by the user, it will be analyzed by the listener trainer to check whether the given query is identifiable or not.

Step 3: After identifying the input, it will split the given query by tokenization and stemming process, which assembles all the harmonized forms of a word in order to rend them down to their root dictionary form or lemma.

Step 4: With the help of Naïve Bayes' algorithm the probability has been identified, which act as the key.

Step 5: This key will hit the database for the output. After identifying the solution, this will wait for sometimes to check whether the identified solution is suitable to the user's query or not by run and wait function in algorithm.

Step 6: At last it will convert those query in voice format by predefined function. This will automatically change the replay in voice format and display the output in both voice and text format in console.

IV. RESULTS AND DISCUSSION

The product is experimented with both the Low sugar based queries and High sugar based queries, and got the output in both sensory and textual method. This product is designed using Naïve Bayes' algorithm to identify the probability which act as the key. The key was corralled by stemming property. This property was abscond the positive and negative arguments from the sentences pattern and hit the database. At last it will exhibit the both textual and sensory responses.

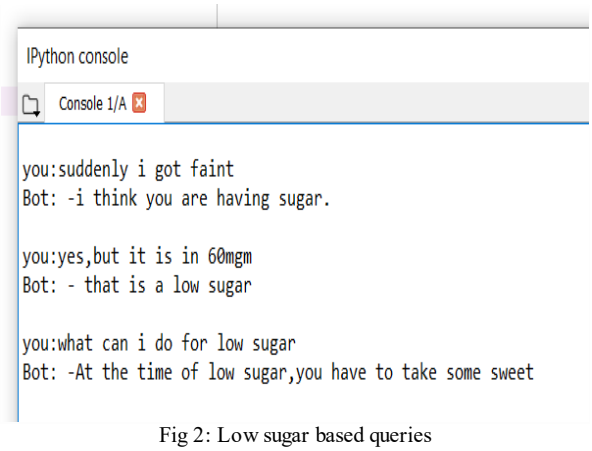


Fig 2: Low sugar based queries

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 06,2022 at 10:17:42 UTC from IEEE Xplore. Restrictions apply.

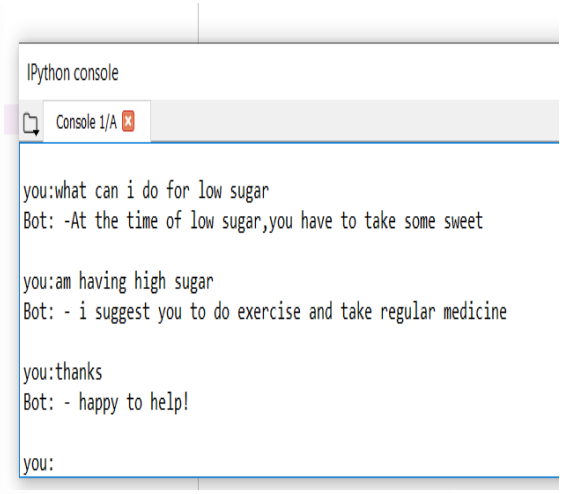


Fig 3: High sugar based queries

V. CONCLUSION

The opious chatbot systems are scrutinized which succeed in practical domains like education, e-commerce as well as health care. By referring those existing system, a voicebot for health care counselling is proposed. This hunk of software helps to identify the commonly infected diseases like hypertension, diabetics, cold, fever and head ache. It describes about the few generic algorithms which have been used in the Bayes algorithm. When contemplating to all these bots, the peculiarity of the proposed system is that can get retort by both sensory and textual methods. This graces the uneducated users to use proficiently. It bears the whole system in patron's cordial manner.

VI.REFERENCES

[1] Luliana Marin, Nicolae Goga "preeclampsia counseling via chatbot for pregnant women"11th International conference on Big Data and Smart Computing,2017
[2] G.Leifer "A Nursing care of women with complications during pregnancy" IEEE Conference on Big Data and Smart Computing. on pp.90-95, Saunders,2014
[3] K.Oh, D.Lee, B.Ko and H.Choi, "A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based Emotional Dialogue Analysis and Sentence Generation", IEEE International Conference on Big Data and Smart Computing.
[4] K.Bala sahila Pandita," chatbot for management using AI, Sentimental analysis" [nov 2017]
[5] Dongkeon Lee,Kyo-Joong Oh and Ho-Jin Choi," A Chatbot feels you-a counseling service using emotional response generation" IEEE International
[6] Agnese Augello, Giovanni Pilato, "An Approach to enhance chatbot semantic, maintainability", IEEE.
[7] Emanuel Haller, Traian Reti, "Chatbot that simulates a historical figure, Analyses the querying technology" IEEE[2013]
[8] A.Sheth, H.Y.Yip and S.Shekharpour,"Extending Patient-Chatbot Experience with Internet-of-Things" IEEE Intelligent System [july-aug 2019].Conference on Education and New Technologies.
[9] A.Folstad, et al P.Brandtzaeg,"New world of HCI Interactions", 2017.
[10] F.T.Spradley,"Pathophysiology of hypertension in preeclampsia", a lesson in Integrative Physiology vol.208, no.3, 2015.
[11] R. Mustafa, et al S. Ahmed, A. Gupta, R. C. Venuto, "A Comprehensive Review of Hypertension in Pregnancy," Journal of Pregnancy [2012].
[12] E.Bartsch,K.E. Medcalf,A.L.Park j," Clinical Risk Factors For Preeclampsia" British Medical Journal
[13] Marin, Goga, "Learning Analytics Software for Medical Students Regarding Pregnancy.
[14] I.Marin,N.Goga,"Treatment Assignment based on Automatic Learning For Pregnant" IEEE
[15] Emanule Haller, Traian Reti, "Designing a chatbot that simulates a historical figure" IEEE.
[16] Revathy.S,B. Parvaathavarthini ,S.Rajathi," Futuristic validation method for rough fuzzy clustering" Indian Journal of Science and Technology 8,no.2(2015)

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 06,2022 at 10:17:42 UTC from IEEE Xplore. Restrictions apply.
View publication stats

Brazilian Journal of Microbiology
https://doi.org/10.1007/s42770-020-00263-8

BIOTECHNOLOGY AND INDUSTRIAL MICROBIOLOGY - RESEARCH PAPER

Green synthesis of gold nanoparticles using extracellular metabolites of fish gut microbes and their antimicrobial properties

T. Rajasekar[1] · K. Karthika[2] · G. Muralitharan[2] · A. Maryshamya[1] · S. Sabarika[2] · S. Anbarasu[1] · K. Revathy[1] · N. Prasannabalaji[3] · S. Kumaran[4]

Received: 11 August 2019 / Accepted: 18 March 2020
© Sociedade Brasileira de Microbiologia 2020

Abstract

In the present study, we synthesis nanoparticles using biosynthesis methods because of the eco-friendly approach. Gold nanoparticles were synthesized using extracellular metabolites of marine bacteria (*Rastrelliger kanagurta, Selachimorpha sp.,* and *Panna microdon*). After the synthesis gold nanoparticles checked their antibacterial and antimycobacterial activities. Here we have few techniques that have been used for characterizing the gold nanoparticles followed by ultraviolet (UV)-visible spectrophotometer analysis, Fourier transform-infrared spectroscopy (FTIR), X-ray diffraction (XRD), scanning electron microscope (SEM), and transmission electron microscopy (TEM). We observed the formation of gold nanoparticles using UV-Vis spectroscopy (UV-Vis). FT-IR spectroscopy results of the extracellular metabolites showed that different characteristic functional groups are responsible for the bioreduction of gold ions. In the recent years, we used zebrafish for an animal model to estimate nanoparticle toxicity and biocompatibility. We tested toxicity of the gold nanoparticle using the zebrafish larvae that are growing exponentially. Sample 1 showed a good antimicrobial activity, and sample 5 showed a good antimycobacterial activity. Based on the UV spectrophotometer, sample 1 is used for further studies. Color change and UV spectrum confirmed gold nanoparticles. Based on the TEM and SEM particles, size was measured and ranged between 80 and 45 nm, and most of the particles are spherical and are in rod shape. XRD result showed the gold nanoparticles with crystalline nature. Toxicity studies in the zebrafish larvae showed that 50 μg ml^{-1} showed less toxicity. Based on the studies, gold nanoparticle has good antibacterial and antimycobacterial activities. The present was concluded that gold nanoparticles have potential biocompatibility and less toxicity. Gold nanoparticles will be used as a drug molecule in pharmaceutical company and biomedicine application.

Keywords Gold nanoparticles · Marine fish · *Mycobacterium tuberculosis*

Introduction

The WHO (world health organization) has stated that AMR (antimicrobial resistance) as one of the "most important threats" to the world health [1]. In the world, around 75,000 deaths per year have been estimated because of the AMR [2]. They form antibiotic resistance for the increased efflux drugs, development of drug degrading enzymes in microorganisms, and formation of biofilm layer which surrounds the bacteria and avoids its exposure to antibiotics [3]. Treating the diseases with a high amount of drugs which leading to high toxicity and more side effects on human and animals. Nowadays, metallic nanoparticle has proven their effectiveness against the multi-drug resistant microbes and synergy with current conventional antibiotics [4]. Nanobiotechnology is a significant and up-and-coming interdisciplinary research field. Nanobiotechnology is a significant and up-and-coming

Responsible Editor: Gisele Monteiro.

✉ G. Muralitharan
microraja09@gmail.com

[1] Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology (Deemed to be University) Jeppiaar Nagar, Rajiv Gandhi Road, Chennai, Tamil Nadu 600 119, India

[2] Departmentof Microbiology, Centre of Excellence in Life Sciences, Bharathidasan University, Tiruchirappalli, Tami Nadu 620 024, India

[3] P. G. & Research Departments of Microbiology, Sri Sankara Arts and Science College (Autonomous), Kanchipuram, Tamil Nadu, India

[4] Department of Biotechnology, Sathyabama Institute of Science and Technology (Deemed to be University) Jeppiaar Nagar, Rajiv Gandhi Road, Chennai, Tamil Nadu 600 119, India

interdisciplinary research field. Since the last decade, Gold nanotechnology used in the various filed because of the potential application in biomedical, electronics, and pharmaceutical fields [5]. In the latest years gold nanoparticle was extensively studied because of the environmental friendly, low-cost and bio-applicable for biosynthesis [6]. Nanoparticles were synthesis from biological sources like bacteria, fungi, actinomycetes, and plants [7]. Nanoparticles were synthesis from extracellular products which are execrated by microorganisms [8]. Nanoparticles were synthesized using intracellular products of microbes. It has difficult to synthesis because more processing steps were required [9]. Nowadays, the green chemistry procedure which uses the microorganisms and plant for synthesis of nanoparticles has received more attention in scientific groups. Nano biotechnology research an emerged as a burgeoning area. Nanoparticles was synthesized using microorganisms to get more potential compounds. In the latest years gold nanoparticle was extensively studied because of the environmental friendly, low-cost and bio-applicable for biosynthesis [6]. In the present study to synthesis the gold nanoparticles using bacterial supernatant, which was isolated from Indian marine fish. Although to evaluate antibacterial and anti-mycobacterial activities and characterizing the gold nanoparticles were carried. We tested toxicity of the compound in zebrafish larvae.

Materials and methods

Chemicals

Gold chloride (AuCl3; Merck, India) was used in this study. We got all the microbiological media and chemicals from HiMedia Laboratories, India. We used ultrapure Milli Q water throughout the study.

Strains used

Serratia marcescens, *Streptococcus mutants*, *Candida albicans*, *Proteus* sp., *Pseudomonas fluorescens*, *Micrococcus luteus*, *Salmonella typhi*, *Escherichia coli*, and *Staphylococcus aureus* were used to test antimicrobial activity, and MTB H37Rv/all-sensitive MTB/MDR MTB was tested for antimycobacterial activity. All the microbial cultures were taken from Centre for Drug Discovery and Development at Sathyabama Institution of Science and Technology, Chennai.

Ethical approval

The International Ethical Committee's recommendation on "Guidelines for the use of fish in research" was followed for fishes while handling (Jenkins et al., 2014) [10].

Fish sample collection

Marine fishes (*Rastrelliger kanagurta*, *Selachimorpha* sp., and *Panna microdo*n) were got from Chennai coastal regions. We transported all the live fish to the lab. We anesthetized the animals with 0.01% benzocaine. Their intestines were excised and then washed with sterile 0.65% saline solution to remove adherent bacteria [11].

Isolation of bacteria from fish gut

We dissected the fish to remove the digestive track in sterile condition. Then, it was homogenized with the help of mortar and pestle along with a little amount of sterile seawater. We diluted 1 ml of fish suspension in sterile seawater. We diluted 1 ml of fish suspension in sterile seawater and taken 0.1 μl of sample was taken from dilution and spread in the nutrient agar (prepared with 100% seawater). The plates were incubated at 30 °C for 48 h. Individual colonies were subcultured and stored for further analysis (Fig. 1).

Biosynthesis of gold nanoparticles using the bacterial supernatant

We inoculated isolated bacterial strain in nutrient broth prepared with seawater incubated at room temperature for 24 h. Followed by supernatant (cell-free extract) collected after centrifugation (at 1800 rpm) was used for the synthesis of gold nanoparticles. A 10 ml of supernatant was added to 90 ml of 0.1 mM aqueous AuCl3 solutions at room temperature for about 24 to 48 h. UV-Vis spectra of the solution monitored the bioreduction of Au^{+} ions in aqueous solution. We recorded UV as an act of time of reaction on a UV-1600 Shimadzu spectrophotometer (400–700 nm) [12].

Antimicrobial activity test

Antibacterial activity of bacterial supernatant using well diffusion method

We seeded Mueller Hinton medium with 24 h culture of bacterial strains (*Serratia marcescens*, *Streptococcus mutans*, *Candida albicans*, *Proteus* sp., *Pseudomonas fluorescens*, *Micrococcus luteus*, *Salmonella typhi*, *Escherichia coli*, and *Staphylococcus aureus*). We cut wells using the well cutter, and we added a 100 μl of the nanoparticle extract. After that we incubated all the plates at 35 °C for 24 h.

Antibacterial activity of gold nanoparticles using well diffusion method

We seeded Mueller Hinton medium with the 24 h culture of bacterial strains. Wells were cut, and 100 μl of the

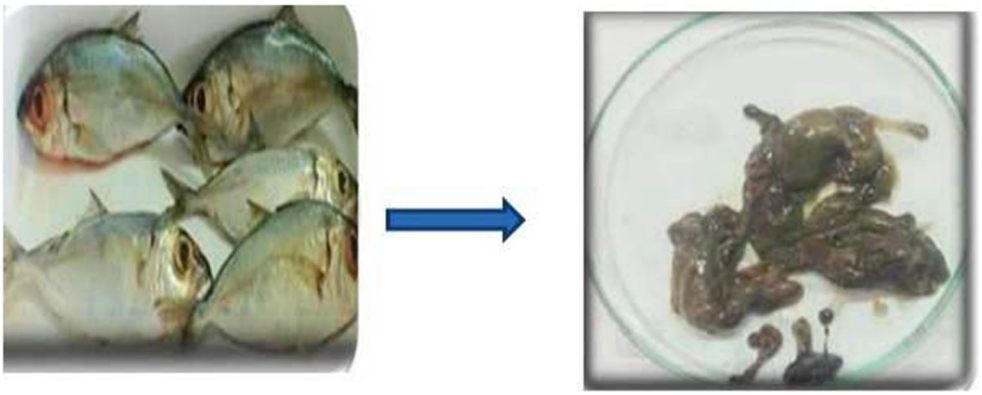


Fig. 1 Fish: dissected to remove the digestive track

nanoparticle extract stock (10 mg/ml) in different dilutions (20, 50, 75, and 100) were added. The plates were then incubated at 37 °C for 24 h. The antibacterial activity was assayed by measuring the diameter of the inhibition zone formed around the well [13].

MIC (minimum inhibitory concentration) test

We seeded Mueller Hinton medium with the 24 h culture of bacterial strains. Wells were cut, and 100 μl of the nanoparticle extract stock (10 mg/ml) in different dilutions (20, 50, 75, and 100) were added. The plates were then incubated at 37 °C for 24 h. The antibacterial activity was assayed by measuring the diameter of the inhibition zone formed around the well [13].

Antimycobacterial assay (LRP)

We screened the antimycobacterial activity of the nanoparticles against *M. tuberculosis* H37Rv and HR-sensitive and HR-resistant *M. tuberculosis* using luciferase reporter phage (LRP) assay [14]. A 400 μl of Middlebrook 7H9 broth was added in cryo vial (control), and each 350 μl of Middlebrook 7H9 broth was aliquoted into two cryo vials and named as Test-500 and Test-250. It added 50 μl of nanoparticles to the test vials at the desired concentration of 500 μg/ml and 250 μg/ml. About 100 μl of *M. tuberculosis* H37Rv suspension was added to all the vials and incubated at 37 °C for 72 h. After incubation, 40 μl of 0.1 M CaCl2 and 50 μl of mycobacteriophage were added to all the vials and further incubated for 4 h. The 100 μl of aliquot from each vial was taken and added with 100 μl of D-luciferin to measure RLU using luminometer. Then, the percentage of reduction in RLU was measured by the formula:

$$\text{Percentage of Reduction in RLU} = \frac{\text{Control RLU} - \text{Test RLU}}{\text{Control RLU}}$$

We performed the above-mentioned LRP assay for HR-sensitive and HR-resistant *M. tuberculosis*.

Characterization of gold-nanoparticles

We use several techniques for characterizing different nanoparticles. Here we have few techniques that have been used for characterizing the gold nanoparticles followed by ultraviolet (UV)-visible spectrophotometer analysis, Fourier transform-infrared spectroscopy (FTIR), X-ray diffraction (XRD), scanning electron microscope (SEM), and transmission electron microscopy (TEM).

Analysis by ultraviolet (UV)-visible spectrophotometer

The formation of AuNPs was determined using a UV-VIS Shimadzu1800 spectrophotometer (igh resolution (1 nm bandwidth) (Shimadzu, Kyoto, Japan), by scanning the samples from300 to 700 nm at a scan speed of 480 mm/min. We performed baseline correction of the spectrophotometer using a blank reference

Analysis by Fourier transform-infrared spectroscopy (FTIR)

Fourier transform-infrared spectroscopy (FTIR, Shimadzu, Japan) was used to name the biomolecules responsible for the reduction of the Au+ ions and capping of the bioreduced gold nanoparticles synthesized by bacteria. To determine the functional groups and their involvement in the synthesis of gold nanoparticles, we carried FTIR analysis out as described earlier [15].

Analysis by X-ray diffraction (XRD)

X-ray diffraction obtained the crystallographic information of gold nanoparticle (XRD) analysis on Bruker D8 ADVANCE diffractometer with CuKα (1.5406 Å) radiation. The XRD was operated at 40 kV and 40 mA at a 2 theta range of 20–80°. Prior to XRD analysis, we centrifuged the gold colloidal solution at 12,000 rpm for

20 min. The pellet was washed and suspended in DW. Sample was freeze-dried and heated at 300 °C for 1 h. It can carry out the XRD analysis [16].

Analysis by scanning electron microscope (SEM)

We used SEM (Scanning electron microscope) to visualize the morphology and size of the nanoparticle. Preparation of slides by adding small drops of nanoparticles suspension and it was left to dry. After the process the slide was analyzed by SEM (JEOL-MODEL 6390 machine). The microscope operated at a speed up voltage of 5–10 kV at different magnification, low vacuum, a spot size 4, and working distances 5–10 mm [17].

Analysis by transmission electron microscopy (TEM)

The morphological analysis of the nanoparticles was done with transmission electron microcopy (TEM). A drop of aqueous gold nanoparticle sample was loaded on carbon-coated copper grid, and it could dry completely for an hour at room temperature. We recorded the TEM micrograph images on a JEOL 1200 EX instrument on carbon-coated copper grids with a speeding up voltage of 80 kV. The clear microscopic views were observed and documented in different ranges of magnifications.

Toxicity study using zebrafish

We have collected zebrafish embryos based on the procedure depicted by Westerfield [18]. Zebrafish embryos were washed with distilled water for 3–4 times and then transferred to petri dish containing egg water, and healthy embryos were chosen at the 8-cell stage. For toxicity studies, we transferred 10 healthy post-hatched zebrafish to the wells of a 24-well plate along with 1 ml of embryo water (60 mg of sea salt/l of ultrapure water). Different concentrations of nanoparticles (5, 10, 25, 50, and100 μg ml^{-1}) were added to the wells and incubated for 72 h at 28.5 °C. We performed tests in duplicate and repeated thrice (60 embryos per concentration). Mortality of the zebrafish was noted after 24, 48, and 72 h. The embryos appeared opaque and white. The dead embryos were degraded soon, because the structures of intact embryos were more visible by 48 hpf which allowed a clear distinction between the dead and alive. We calculate the mortality rate. At the end of the incubation period, we photographed the embryos using a phase-contrast inverted microscope Olympus CKX41.

Results

Isolation of fish guts bacteria and antibacterial activity of the cell-free extract

A total of 20 (CFU/ml) colonies were isolated from three samples, from that isolates five different morphologically colonies were taken for the antimicrobial activity. The fish gut isolated bacterial supernatant was tested against human bacterial pathogens by agar well diffusion method and the plated were incubated for about 24 h. Nutrient broth is loaded as control. Fish gut isolates supernatant were tested against the bacterial pathogens and exhibited excellent antibacterial activity. Out of five bacteria, three bacteria showed good activity against human pathogen. Sample 1 showed good activities against *Staphylococcus aureus*, *Micrococcus luteus*, *Candida albicans*, *Streptococcus mutans*, *Pseudomonas fluorescens*, *Proteus* sp., and *Serratia marcescens.* Sample 3 showed *Staphylococcus aureus*, *Streptococcus mutans*, *Micrococcus luteus*, *Candida albicans*, *Proteus* sp., *Pseudomonas fluorescens*, *Serratia marcescens*, and *Escherichia coli*. Sample 5 showed potential activities against *Micrococcus luteus* followed by *Staphylococcus aureus*, *Candida albicans*, *Streptococcus mutans*, *Pseudomonas fluorescens*, and *Serratia marcescens.* Whereas sample 2 and 4 did not show any activity against human pathogens.

Synthesis gold nanoparticles

A potential cultures of sample 1 and 5 bacterial supernatant were taken for the synthesis of gold nanopartles. Change of the color in the microbial supernatants was from yellow color to deep purple after the incubation of 24 h after 48 h intensity of color was increased. In the control, color did not change (Figs. 2).

Antimycobacterial activity of the gold nanoparticle

Gold nanoparticles showed a good antimycobacterial activity. Prospective cultures (sample 1 and 5) were taken for the synthesis of gold nanoparticles, and anti-mycobacterial assay was screened utilizing luciferase reporter phage (LRP) test against the mycobacterium. Gold nanoparticles synthesized from sample 1 showed 61.47%, 82.52%, and 75.5% RLU reduction against *M. tuberculosis* H37Rv, HR-sensitive *M. tuberculosis*, and HR-resistant *M. tuberculosis*, respectively. Whereas sample 5 confirmed inhibition against *M. tuberculosis* H37Rv (91.13%), HR-sensitive *M. tuberculosis* (86.72%), 276 and HR-resistant *M. tuberculosis* (77.10%) RLU (Figure. 12). Sample 3 did not shown any activity against *M. tuberculosis*.

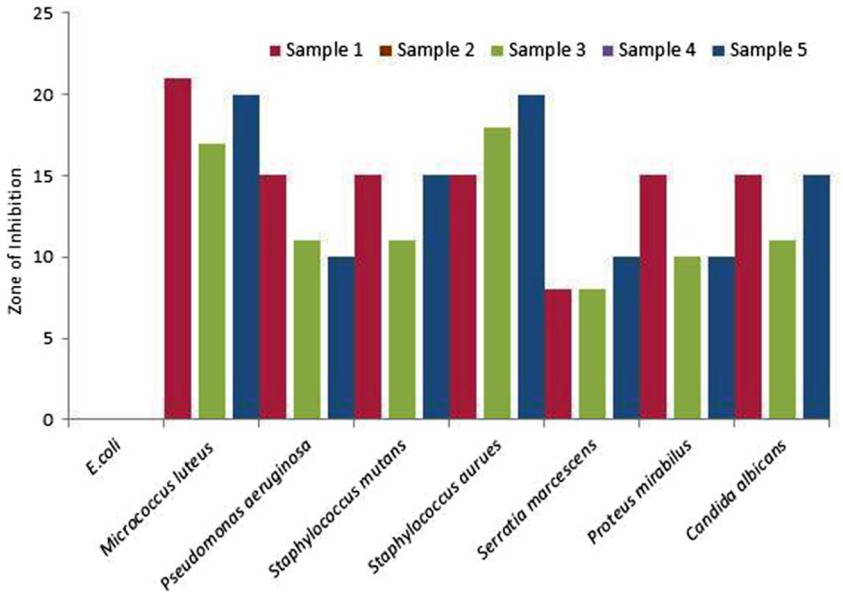


Fig. 2 Antibacterial activity of bacterial supernatant

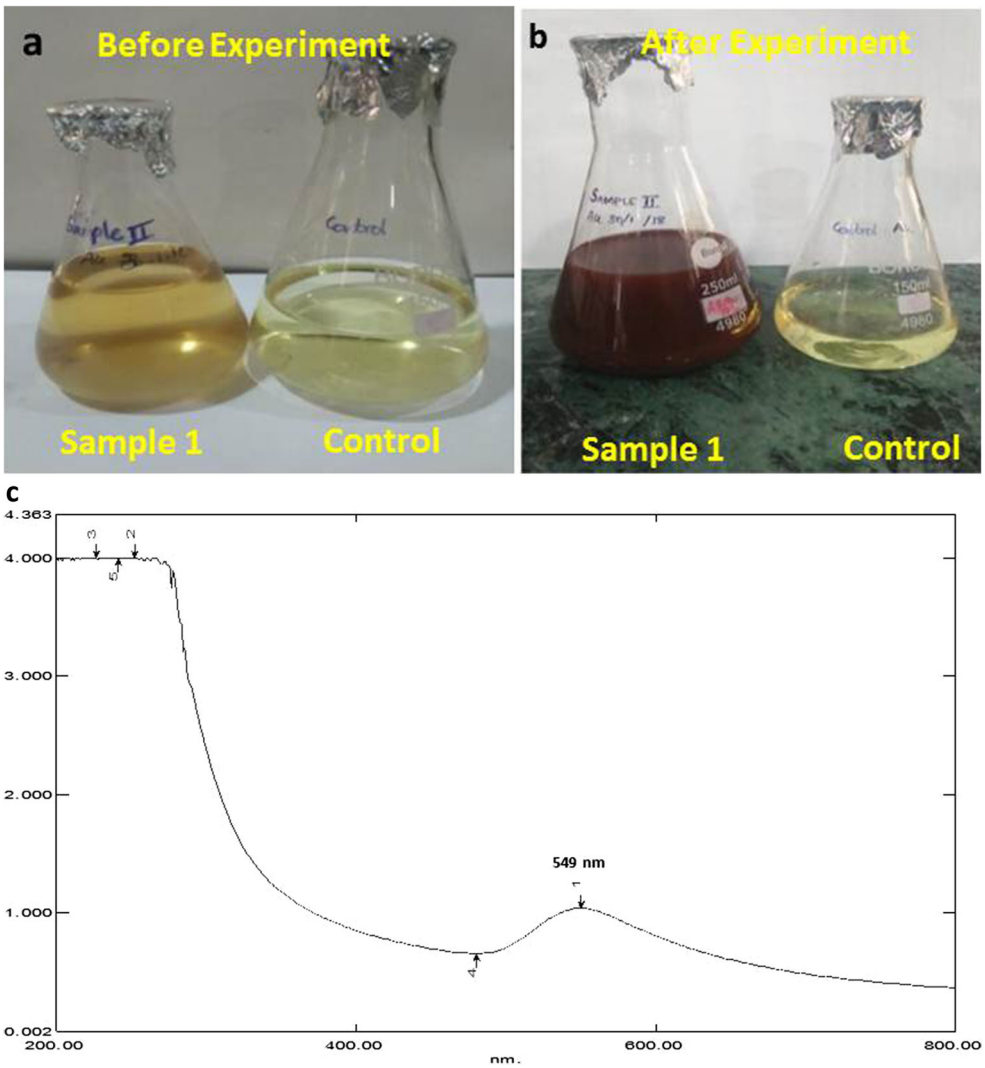


Fig. 3 **a** Biosynthesis of Gold Nanoparticles using bacterial supernatant (Sample 1). **b** UV-visible spectrophotometry of Gold Nanoparticles synthesized from bacterial supernatant (Sample 1)

Characterization of gold nanoparticles

Ultraviolet (UV)-visible spectrophotometer analysis

The formation of gold nanoparticles in bacterial supernatant was confirmed by UV- visible spectrophotometer. Optical densities of the gold nanoparticles were taken from 200 to 800 nm. The solution showed maximum a plasma resonance absorbance at 549 nm and 552 nm, respectively (Figs. 3b and 4b). Based on the antibacterial and antimycobacterial activity of sample 1 synthesized nanoparticles was taken for further characterization.

Analysis by scanning electron microscope (SEM)

The results of the sample1 through scanning electron microscope are showed in Fig. 3a. The result showed that nanoparticles of the seaweed were spherical, and size is less than 80–45 nm in size.

Analysis by transmission electron microscopy (TEM)

TEM (Transmission electron microscopy) was used to characterize the gold nanoparticle synthesized from the cell free extract of the potential strain (sample 1). Characterization of the gold nanoparticles is given in 4b. Most of the particles are spherical, and some particles are in rod shape. It was further analyzed by FTIR and XRD.

Analysis by Fourier transform-infrared spectroscopy (FTIR)

FT-IR investigation shows that different characteristic functional groups in turn are responsible for the bioreduction of gold ions by extracellular metabolites. FTIR analysis was performed to determine the functional group of these gold nanoparticles (Table 1). Sample 1 of the FTRI spectrum of the gold nanoparticles are showed in Fig. 6.

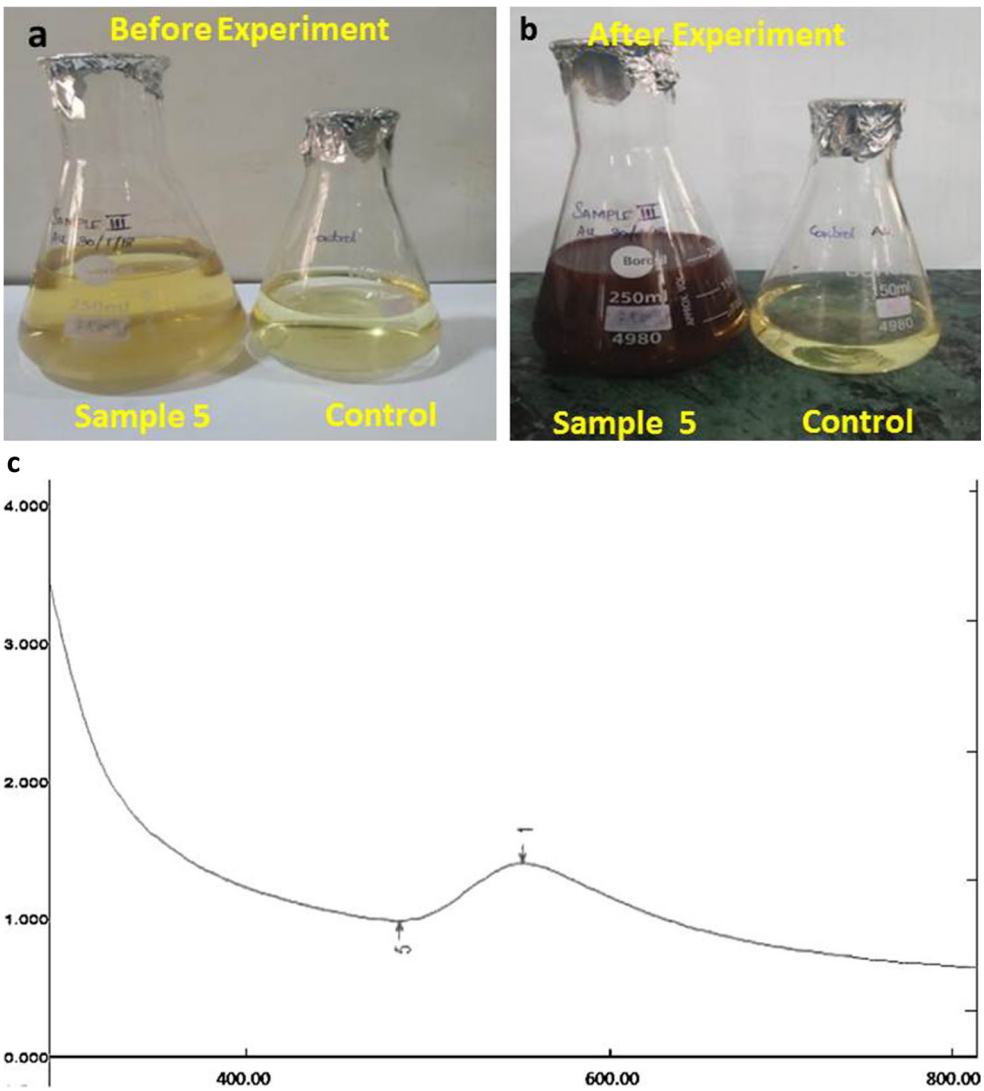


Fig. 4 **a** Biosynthesis of Gold Nanoparticles using bacterial supernatant (Sample 5). **b** UV-visible spectrophotometry of Gold Nanoparticles synthesized from bacterial supernatant (Sample 5)

Table 1 Fourier transform alkyl-infrared spectroscopy spectrum of the gold nanoparticles

Wave number (cm^{-1})	Functional groups	Bond
3888.03	Alcohol	OH
3395.99	Primary amine	NH
2921.93	Alkyl	CH
1582.69	Aromatic C═C	C═C
1313.77	Phenol	OH

Analysis by X-ray diffraction (XRD)

Crystallite nature of $AuCl_3$ was analyzed using XRD. XRD were used to analyze the crystalline structure of the compound. It is necessary to characterize the gold nanoparticle structure. Figure 7 showed the XRD pattern of the produced gold nanoparticles. The peaks appeared at 32°,46°, 57°, and 76° in the 2 range of 0°–80°, parallel to the characteristic diffraction of the (111), (200), 220), (311), and (222). Therefore, XRD showed that $AuCl_3$ formed using cell-free extract of bacterial sample 1 was basically crystalline in nature (Fig 8).

Antibacterial activities of gold nanoparticles

Antibacterial activities of the gold nanoparticles (sample 1) were tested against different pathogens. Maximum inhibition is found in the *E. coli*, followed by fungi Candida albicans, Bacillus cereus, Micrococcus luteus, and *Serratia marcescens*.

Toxicity study using zebrafish

The acute toxicity of the $AuCl_3$ in vivo zebrafish embryo is shown in Fig. 9. The exposure of the nanoparticles (sample 1) at different concentration to zebrafish was 5, 10, 25, 50, and 100μgml^{-1}. Synthesized Gold nanoparticles at the 100 μg ml^{-1} concentration showed high toxicity to zebrafish larvae. Toxicity studies increased the concentration of nanoparticles showed a less survival rate of embryos. After 72 h, all the experiments show the mortality. There is no mortality in the control groups

Discussion

In the present work was emphasized the cell-free bacterial supernatant was used for the synthesis of gold nanoparticles and evaluated their antibacterial and antimycobacterial

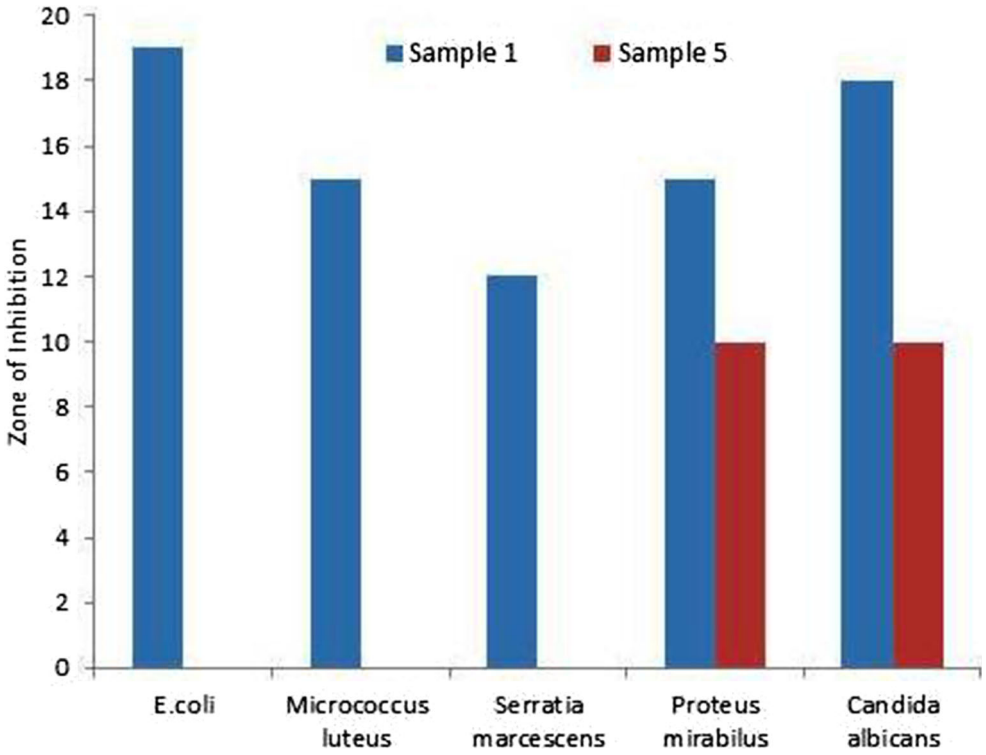


Fig. 5 Antibacterial activities of the Gold Nanoparticle of bacterial supernatant (sample 1 and sample 5)

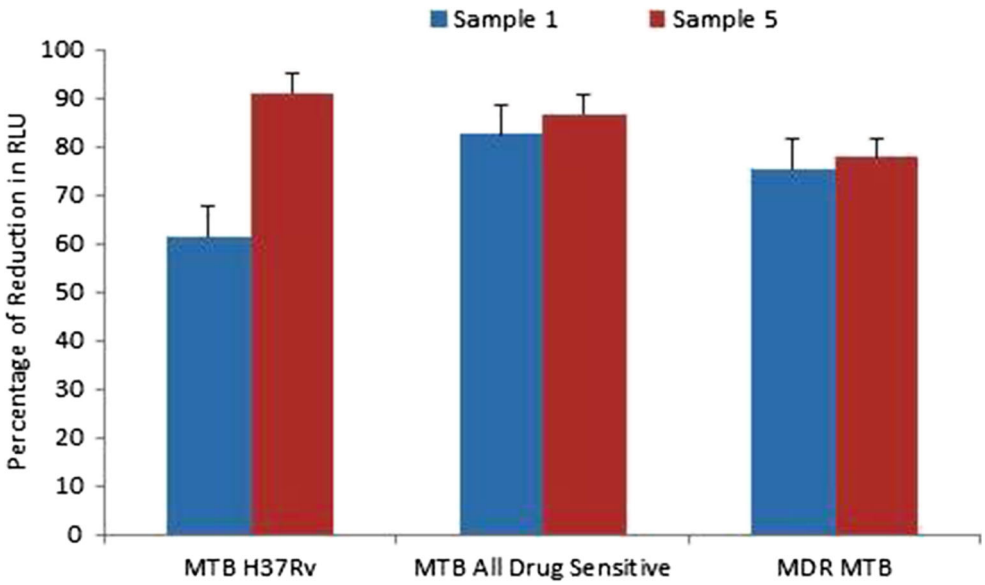


Fig. 6 Anti-mycobacterial activity against Gold Nanoparticle Seaweeds (sample 1 and sample 5)

activities against human pathogens. The significance of the interaction between microbes and metals has been altering their physical and chemical state, also affecting the growth of the microorganism. The most conformation of the bioreduction of the metallic ions changes in their native color. Luo et al. [19] reported that in the ancient time, many number of metals have been used in as antibacterial agent. Gold nanoparticles were identified by color change from light yellow to dark purple color after the incubation of the bacterial extracts samples with $AuCl_3$. Current investigation was to synthesis of gold nanoparticles from cell free supernatant of potential strains isolated from fish gut.

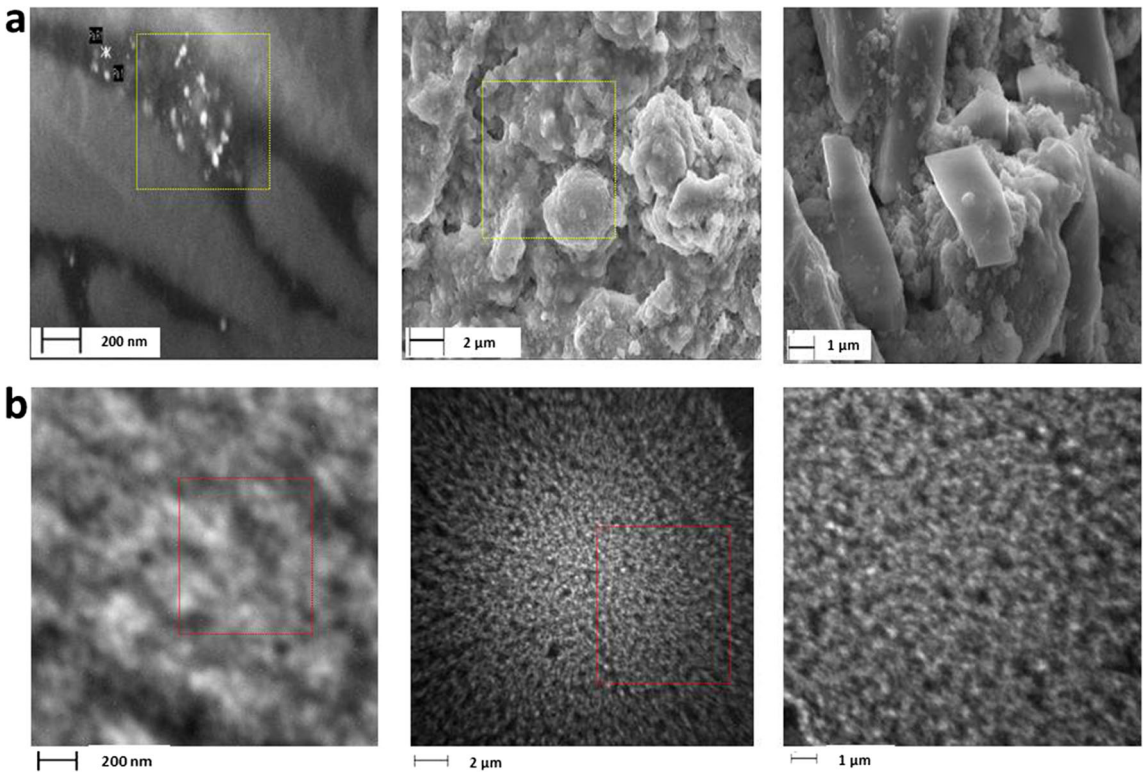


Fig. 7 **a** Analysis of Gold Nanoparticles using SEM of sample 1. **b** Analysis of gold nanoparticles using TEM of sample 1

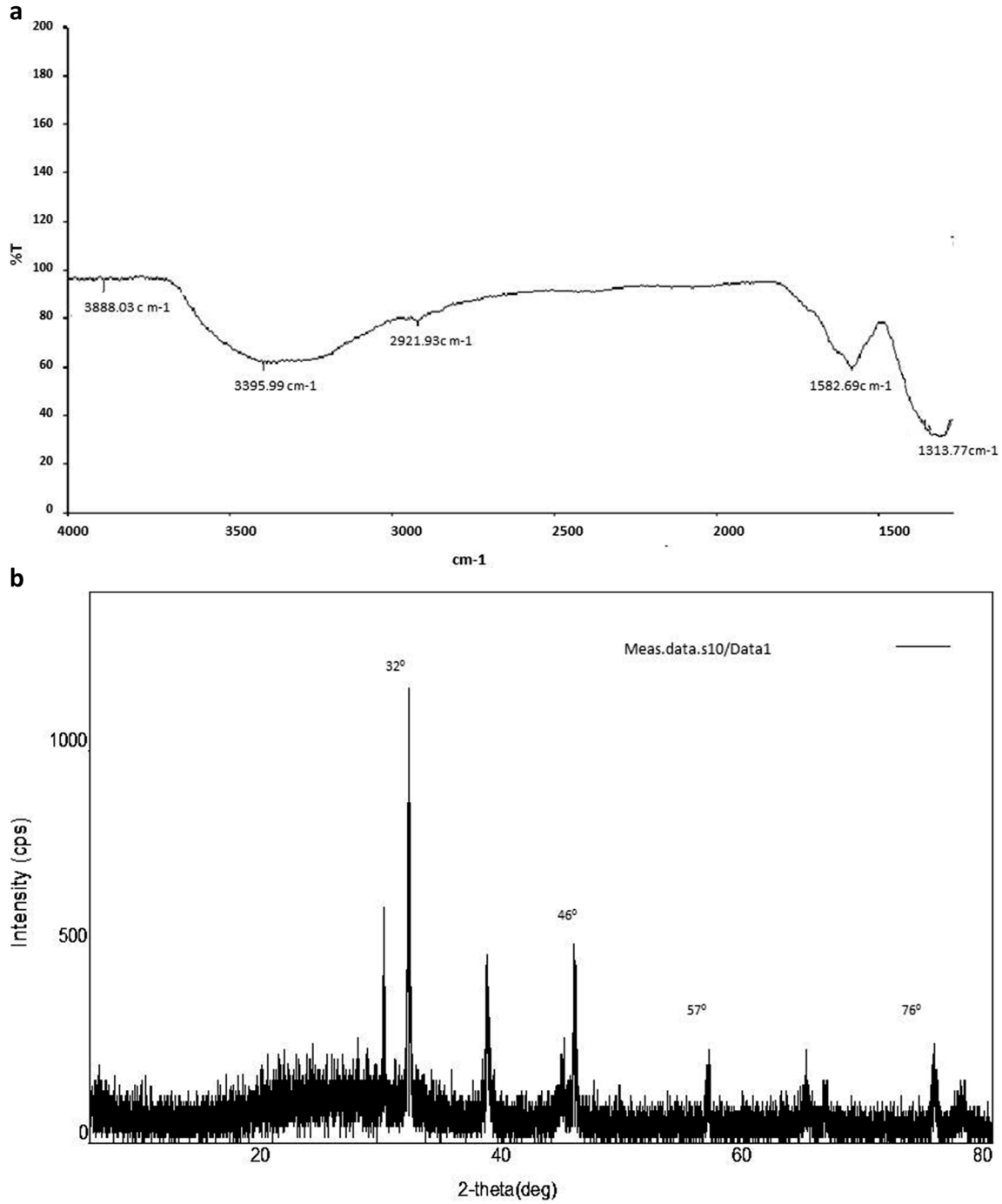


Fig. 8 **a** Analysis of Gold Nanoparticles using FTIR spectrum of sample 1. **b** Analysis of gold nanoparticles using X-Ray diffraction of sample 1

Khan et al. [20] reported that gold nanoparticles bind with organic molecules. It is used as a vaccine carrier and therapeutic agents which has increased the drug effectiveness and destruction of pathogens. Nanotechnology is very important because of the small physical size and nanoscale size, and some have toxic effects [21]. In recent trends, different types of nanoparticles are synthesis by utilizing various sources. By using natural resources, the possibility of producing nanoparticles with low toxicity in-vivo models [22]. Gold nanoparticles were widely used in a different research area like nanomedicine. In the present study is to evaluate the antimycobacterial and antibacterial activities of gold nanoparticles and it has been tested against various microorganism like *Mycobacterium tuberculosis* and *Serratia marcescens, Staphylococcus epidermidis, Candida albicans, Proteus sp.,Pseudomonas fluorescens, Micrococcus luteus, Salmonella typhi, Escherichia coli,* and *Staphylococcus aureus*. Solano-Umana et al. [23] suggested that these gold

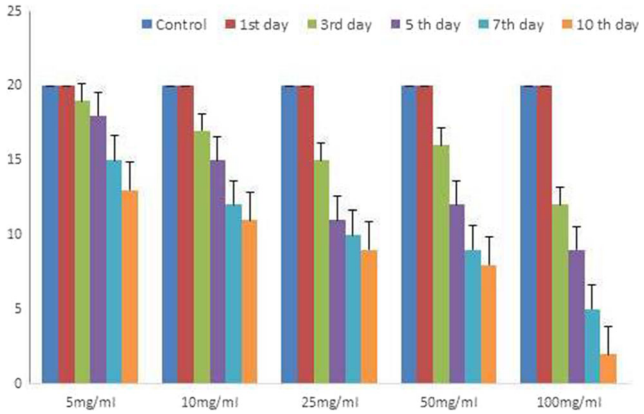


Fig. 9 Toxicity study of Gold nanoparticles of Sample 1 using zebra fish embryo

nanoparticles are combined with drugs to reduce the side effects, increasing the quality of the patients. It has a strong affinity with biomolecules like pathogens, antibodies, peptides, oligonucleotides, and proteins.

The confirmation of the formation of gold nanoparticles in the microbial supernatants was by ultraviolet (UV)–visible spectrophotometer analysis. Optical densities of the gold nanoparticles were taken from 200 to 800 nm. The solution showed a plasma resonance absorbance at 549 nm. Correspondingly Nejad et al. [24] reported that they experimentally observed UV spectra of 545 nm for the gold nanoparticles from the marine bacteria *Enterococcus* sp. Biradar et al. [25] reported that synthesis gold nanoparticles using bacteria *Klebsiella pneumoniae* produced surface plasmon resonance values in the range of 400–700 nm. Zhang et al. [26] suggested the color change of the gold nanoparticles based on the principle of surface plasmon resonance (SPR). During the color change, the size of the particles was expanded; in that circumstance, gold nanoparticle color was changed from dark red to dim purple. Variables of the color change are due to the SPR that they display in the visible region of the electromagnetic spectrum. The color change of the absorbance was measured using UV-visible spectroscope.

Vanaja et al. [27] suggested that carboxylate and amines functional groups have served as reducing agent in the synthesis of silver and gold nanoparticles and also involved in the reduction and stabilizing mechanisms. Bioreduction of the gold ions to gold nanoparticles and stabilizing mechanisms was confirmed by FTIR. The peak was observed at 3289.34 matches with intramolecular H-bond , –OH of alcohols, followed by peak 2862.76 which corresponds to C–H stretching of alkanes; peak 2129.99 compares to isonitriles for alkynes, peak 1534.08 relates to secondary amine salts of amines, peak 1458.03 corresponds to C–H deformations of alkanes, and peak 1345.45 corresponds to N–O symmetric. Rajasree et al. [28] reported that gold nanoparticles are synthesized from the bacterium *Pseudomonas fluorescens*. SEM analysis showed that gold nanoparticles were very small and almost spherical. nanoparticles were spherical. This result was correlated with our studies. Our result showed that nanoparticles of the seaweed were spherical, and size is less than 90–45 nm. Transmission electron microscopy images were obtained from gold nanoparticles of the cell-free extract of the microorganisms. Most of the particles are spherical, and some particles are in rod shape. Correa-Llanten et al. [29] reported that TEM was used to verify the intracellular particles size of gold nanoparticles. The shape of the particles is quasi hexagonal, and size ranges from 50 to 5 nm. Most of the particle size was 10–20 nm. Kikuchi et al. [30] reported the XRD spectrum of the gold nanoparticle synthesis using *Lactobacillus casei*. In the XRD, results showed that several peaks corresponded to the gold crystals with face-centered cubic structure. In these results indicated that *L. casei* produced the gold nanoparticles. Similar results was seen in our results. Accordingly, XRD showed that $AuCl_3$ formed using cell-free extract of bacterial sample was basically crystalline in nature. Xia et al. [31] reported that toxicity assessment is the important factor for nanoparticles in the event of an environment. In our present study also support that gold nanoparticle showed a less toxicity against zebrafish embryo. Similarly, Bakri et al. [32] suggested that Ag-NP causing high toxic and Au-NP is considered as non-toxic in vitro. Zhu et al. [33] reported that metal oxide and metal nanoparticles caused harmful effects in vivo and in vitro. These above statements were supported by AshaRani et al. (2009) [34]. Bar-Ilan et al. (2009) [35] also supported that gold nanoparticles were non-toxic in zebrafish embryos.

Conclusion

In the most recent decades, zebrafish larvae are utilized for toxicity examinations because of their biocompatibility, minimal effort, and simple upkeep capacities. Based on our studies, gold nanoparticle has good antibacterial and antimycobacterial activities. In the present study concluded that gold nanoparticles possess potential biocompatibility and less toxicity, hence it can be used in pharma and biomedicine application.

Funding information The authors are grateful to the Science and Engineering Research Board (SERB-Early Career Research Award ECR/2015/000460), Government of India, New Delhi, India for the financial support and also thankful to the Management of Sathyabama Institute of Science and Technology (Deemed University) Chennai, Tamil Nadu, India, for providing the Infrastructure facility.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Davis MH (2018) Empathy: a social psychological approach. Routledge
2. Betts JW, Hornsey M, La Ragione RM (2018) Novel antibacterial: alternatives to traditional antibiotics. AdvMicrob Physiol 73:123–169
3. Blecher K, Nasir A, Friedman A (2011) The growing role of nanotechnology in combating infectious disease. Virulence. 2(5):395–401
4. Singh P, Chauhan NR, Rajesha S (2019) Influence of cobalt, iron and copper on microstructure and mechanical properties of alumina/SiC nano-ceramic composite. Mater Res Express 6(6): 065027
5. Vazquez E, Villaverde A (2010) Engineering building blocks for self-assembling protein nanoparticles. Microb Cell Factories 9(1): 101
6. Camas M, Celik F, Sazak Camas A, Ozalp HB (2019) Biosynthesis of gold nanoparticles using marine bacteria and box–Behnken design optimization. Part Sci Technol 37(1):31–38
7. Jayaramudu T, Raghavendra GM, Varaprasad K, Sadiku R, Raju KM (2013) Carbohydr Polym 92:2193
8. Nangia Y, Wangoo N, Sharma S, Wu J, Dravid V (2009) J ApplPhys Lett 94:233901
9. Huang X, Neretina S, EI-Sayed MA (2009) Adv Mater 21:4880
10. Jenkins JA, Bart HL Jr, Bowker JD, Bowser PR, MacMillan JR, Nickum JG, Rachlin JW, Rose JD, Sorensen PW, Warkentine BE, Whitledge GW (2014) Guidelines for use of fishes in research-revised and expanded. Fisheries. 39(9):415–416
11. Yang L, Wang Y, Zhang Z, He S (2014) Comprehensive transcriptome analysis reveals accelerated genic evolution in a Tibet fish, Gymnodiptychuspachycheilus. Genome Biol Evol 7(1):251–261
12. Agnihotri M, Joshi S, Kumar AR, Zinjarde S, Kulkarni S (2009) Biosynthesis of gold nanoparticles by the tropical marine yeast Yarrowialipolytica NCIM 3589. Mater Lett 63(15):1231–1234
13. National Committee for Clinical Laboratory Standards (1993) Performance standards for antimicrobial disk susceptibility tests. Approved standard. NCCLS document M2-A5. Wayne, Pa: National Committee for Clinical Laboratory Standards
14. Sivaraj A, Kumar V, Sunder R et al (2019) Commercial yeast extracts mediated green synthesis of silver chloride nanoparticles and their anti-mycobacterial Activity. J Clust Sci. https://doi.org/10.1007/s10876-019-01626-4
15. Bankar AV, Kumar AR, Zinjarde SS (2009) Removal of chromium (VI) ions from aqueous solution by adsorption onto two marine isolates of Yarrowialipolytica. J Hazard Mater 170(1):487–494
16. Pourali P, Baserisalehi M, Afsharnezhad S, Behravan J, Alavi H, Hosseini A (2012) Biological synthesis of silver and gold nanoparticles by bacteria in different temperatures (37°C and 50 C). J Pure Appl Microbiol 6:757–763
17. Caroling G, Tiwari SK, Ranjitham AM, Suja R (2013) Biosynthesis of silver nanoparticles using aqueous broccoli extract-characterization and study of antimicrobial, cytotoxic effects. Asian J Pharm ClinRes 6(4):165–172
18. Westerfield M (2000) The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio). Oregon press, Eugene
19. Luo Z, Zheng K, Xie J (2014) Engineering ultrasmall water-soluble gold and silver nanoclusters for biomedical applications. Chem Commun 50(40):5143–5155
20. Khan AK, Rashid R, Murtaza G, Zahra A (2014) Gold nanoparticles: synthesis and applications in drug delivery. Trop J Pharm Res 13(7):1169–1177
21. Borm PJA, Müller-Schulte D (2006) Nanoparticles in drug delivery and environmental exposure: same size, same risks? Nanomedicine:235–249
22. Joseph BM (2015) Microwave assisted facile green synthesis of silver and gold nanocatalysts using the leaf extract of Aervalanata. Spectrochim Acta A Mol Biomol Spectrosc 136:1371–1379
23. Solano-Umaña V, Vega-Baudrit RJ (2015) Gold and silver nanotechology on medicine. J ChemBiochem 3(1):21–33
24. SoltaniNejad M, ShahidiBonjar GH, Khaleghi N (2015) Biosynthesis of gold nanoparticles using streptomycesfulvissimus isolate. Nanomedicine J 2(2):153–159
25. Biradar L (2012) Isolation and screening of gold nanoparticles by microbes. World J Sci Technol:20–22
26. Zhang X, Qu Y, Shen W, Wang J, Li H, Zhang Z, Li S, Zhou J (2016) Biogenic synthesis of gold nanoparticles by yeast Magnusiomycesingens LH-F1 for catalytic reduction of nitrophenols. Colloids Surf A Physicochem Eng Asp 497:280–285
27. Vanaja M, Paulkumar K, Gnanajobitha G, Rajeshkumar S, Malarkodi C (2014) Annadurai G (2014) herbal plant synthesis of antibacterial silver nanoparticles by *Solanumtrilobatum* and its characterization. Int J Met
28. Rajasree SR, Suman TY (2012) Extracellular biosynthesis of gold nanoparticles using a gram negative bacterium Pseudomonas fluorescens. Asian Pacific Journal of Tropical Disease 2:S796–S799
29. Correa-Llanten DN, Muñoz-Ibacache SA, Castro ME, Muñoz PA, Blamey JM (2013) Gold nanoparticles synthesized by Geobacillus sp. strain ID17 a thermophilic bacterium isolated from Deception Island, Antarctica. Microb Cell Fact 12(1):75
30. Kikuchi F, Kato Y, Furihata K, Kogure T, Imura Y, Yoshimura E, Suzuki M (2016) Formation of gold nanoparticles by glycolipids of Lactobacillus casei. Sci Rep 6:34626
31. Xia T, Li N, Nel AE (2009) Potential health impact of nanoparticles. Annu Rev Public Health 30:137–150
32. Bakri SJ, Pulido JS, Mukherjee P, Marler RJ, Mukhopadhyay D (2008) Absence of histologic retinal toxicity of intravitrealnanogold in a rabbit model. Retina. 28:147–149
33. Zhu X, Zhu L, Duan Z, Qi R, Li Y, Lang Y (2008) Comparative toxicity of several metal oxide nanoparticle aqueous suspensions to zebrafish (Daniorerio) early developmental stage. J Environ Sci Health Tox Haz Subst Environ (Eng) 43:278–284
34. AshaRani PV, Low Kah MG, Hande MP, Valiyaveettil S (2009) Cytotoxicity and genotoxicity of silver nanoparticles in human cells. ACS Nano 3:279–290
35. Bar-Ilan O, Albrecht RM, Fako VE, Furgeson DY (2009) Toxicity assessments of multisized gold and silver nanoparticles in zebrafish embryos. Small 5:1897–1910

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-019-7470-2

Framework for Social Media Analytics based on Multi-Criteria Decision Making (MCDM) model

Muruganantham A.[1] · G. Meera Gandhi[1]

Received: 20 October 2018 / Revised: 26 January 2019 / Accepted: 7 March 2019
Published online: 15 April 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

The Social, Mobile, Analytics and Cloud (SMAC) explosion in recent times changed the way the customers look at and collaborate businesses through large world of data or information, often described as "Big Data". With over 1590 million active users, social media such as Facebook, Twitter, WhatsApp, Instagram, LinkedIn etc. send or receive messages or post or access new content every day. Businesses or enterprises understand and extract useful insights from social media platform and transforming it into useful information or knowledge along with their enterprise business data for strategic decision making. A framework for Social Media Analytics based on Multi-Criteria Decision Making (MCDM) model is proposed for social media data and our comprehensive study on large-scale twitter dataset experiment explains how MCDM (TOPSIS) method outperforms against the standard centrality methods using predicted Spreading Size. Two well-known metrics such as MAE (Mean Absolute Error), and RMSE (Root Mean Square Error) are applied to measure prediction accuracy of our MCDM based method against standard methods. TOPSIS (MCDM based) experiences least error accuracy of 9% in MAE and 16% in RMSE than the standard methods to prove that the proposed approach works better than the standard methods.

Keywords Social media analytics · Centrality measures · Spreading size · MAE · RMSE · Twitter data set

✉ Muruganantham A.
amuruganantham@gmail.com

G. Meera Gandhi
drmeeragandhii@gmail.com

[1] Faculty of Computer Science & Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

1 Introduction

In the era of SMAC, social media plays a vital role in our society and provides a platform to communicate, collaborate and share thoughts. Social media is ubiquitous and important for social networking and content sharing [36]. Social media data such as consumer opinions, experiences and sentiments on various product and services helps to improve value of business organizations [20]. Collecting data from social media sites and evaluating that data to make business decisions is called Social Media Analytics (SMA). SMA has been gaining importantance in recent times thanks to 1.5 billion social media users in Facebook, Instagram, Twitter, LinkedIn etc.. Users in these social media networks produce data through blog posts, likes, and comments that are collated and managed through a Big Data platform such as Apache Hadoop, IBM Cloudera. Business enterprises are expected to understand their customers, competitors or partners holistically with insight trends from both social media platforms and business data platforms such as ERP to make well informed business decisions. Though there are many studies available on SMA, no systematic based study is available on how SMA provides insight into business organizations. Our proposed Mutli-Criteria Decision Making (MCDM) based framework helps business organizations effectively leverage on social media insights. Our experiment results on a social media platform dataset were compared and understood how effectively works better than the standard techniques through well-known metrics - MAE (Mean Absolute Error), and RMSE (Root Mean Square Error).

The rest of the paper is organized as follows: The related work is discussed in Section 2, Section 3 explains MCDM model. The proposed framework based on MCDM is discussed in Section 4, the experience results explained through dataset in Section 5 and the performance accuracy is compared in Section 6 and finally findings are summarized as conclusion.

2 Related works

Social media platforms disturbed the way business organizations engage with customers, and their services. It influences customer likes, dislikes behaviors and buying decisions [17]. Social media platforms become powerful marketing channel to increase customer awareness on business product and its service offerings [19]. SMA operations such as recommending users and communities [22], Influence users, and top trending topics identification are popular and it has been used in brand awareness, product or service development, and politics. Social media platform contents are primarily used to understand how people responds like like, dislike, pattern of buying/selling etc. Social media platform helps business organizations to monitor and determines customer interests, sentiments, trending topics and influencers [9]. The other application to enterprises are viral marketing [10, 18, 28], tweets recommendations [31, 32], trust or information propagation [13, 14], affiliation recommendations [33, 38], sentiment analysis and opinion mining [27].

Number of research methods to understand influential users or spreaders [23, 24] on a social media platform such as Twitter indicates that, while a large amount of followers seem to be present among influential user, predictions of a particular influential user is unreliable [2]. Influential user meaning is different from one social network to another. For example, an influence user on Twitter social media is defined by retweets or

mentions [7, 12, 30] whereas the number of likes and comments defines the influence users on Facebook. The challenges related to security risk in the social media or cloud environment using NLP and other techniques [40] were discussed.

Many studies tried to understand user influence using Page Rank Centrality [26, 29] or Degree Centrality [6, 26]. While there are many methods in SMA popular, according to our review, there is little research on using Multi-Criteria Decision Making models for influence user identification in online social media platforms. Also, there is not much systematic approach to determine how SMA brings value and competitive benefits to business organizations. Our proposed SMA based framework addresses these challenges using MCDM, which is elaborated in the next section.

3 Multi-criteria decision making model (MCDM)

Multi Criteria Decision Making (MCDM) methods solve planning and decision making problems contain multiple criteria. MCDM helps decision makers to determins influence users [25] with multiple features to be solved. MCDM is emerged as a popular model and has a wide application domains [21] to find the best solution to choose the best alternative. Various MCDM methods in Table 1 such as ELECTRE [41], PROMETHEE [3], AHP [1, 5], Pugh, TOPSIS [4] etc. are emerged in the recent past and it has its own advantages and disadvantages. Vagueness and Uncertainty is combined with TOPSIS formed Fuzzy TOPSIS to provide solution with incompelete quantitative information [15]. MCDM methods are classified based on the type of data, the number of decision makes invovled. Some MCDM methods are suitable for specific situations, while other applications avoid using certain MCDM methods altogether [39].

4 Proposed framework for social media analytics based on multi-criteria decision making (MCDM) model

Our propsed SMA framework shown in Fig. 1 is described using Mutli-Criteria Decision Making (MCDM) model extends our previous work [11]. The customers data

Table 1 Comparions of popular MCDM featres

Method	Merits	Demertis
PROMETHEE	Simple to use	Unclear on weights assignments
ELECTRE	Uses Outranking	Difficult to explain process and outcome; Long time for processing
AHP	Simple to use; scalable; sutiable to different sized problems due to its hierarchical structure	Irregular Ranking; Large number of comparison are required
Pugh	Decision making is subjective i.,e., opinions about one alternative versus another can be made more objective	Expertise choose selection criteria
TOPSIS	Simple to use and program, Same stps follwed for any problem size.	Complex to keep consistency of judgment

from the enterprise and their social data from social media such as Facebook, Twitter are stored in an integrated Database through Web Services (WS). Web Services are used to integrate applications, which are running on different architecture with different computer languages over Internet-based environment [35]. Web Services uses standards such as XML(eXtended Markup Language), SOAP (Service Oriented Applicatoin Protocol), UDDI (Universal Discovery, Descriptoin and Integration),WSDL (Web Service Descriptoin Language) and HTTP (Hyper Text Transfer Protocol). Data from social media, which is a combination of semi-structured, un-structured form and is a huge size in terms of tera or peta bytes are processed with help of open source Big Data platform- Apache Hadoop. Apache Hadoop framework uses distributed storages, process high amount of data on commodity hardware through Hadoop Distributed File System (HDFS) and MapReduce. Hadoop framework is much faster and cheaper than traditional systems [37]. The extracted results from Apache Hadoop framework is further processed by MCDM based social analytics and intelligence algorithm as explained in Algorirthm 1. The output generated or identifed from the algorithm 1 can be used for enterprise decision making.

Algorithm 1 MCDM based social analytics and intelligence

Input: Social media network data set
Output: MCDM based classified data

Step 1: Select and represent social data using graph network G(V,E) where V describes all users posted at mininum one hashtag or URL and E represents the connection between followers or between friends as shown in Fig. 2.

Step 2: Compute centrality score as shown in Fig. 3 from degree, closeness centrality, betweeneess centrality, eigen centrality and Page Rank

Step 3: compute MCDM based score using the standard centralities scores generated in step 2 as multiple criteria

Using the features available in Gephi, centrality measures such as Degree, Betweenness Centrality, Eigen Vector Centrality and Page Rank are computed as shown in Fig. 3. The open

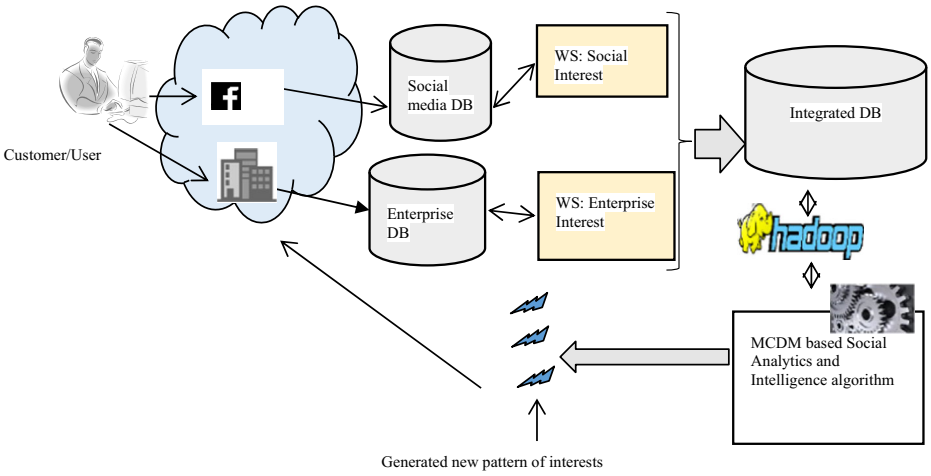


Fig. 1 SMA based architecture framework using MCDM model

source tool R is used to compute MCDM based score using TOPSIS method described in Algorithm 2.

Algorithm 2 MCDM score computation

Input: Centrality methods data on social network in .csv format
Output: MCDM score (using TOPSIS) matrix

Step 1: Read and load criteria data from Degree, Eigen Vector, Page Rank and TOPSIS (MCDM based) as performance table
Step 2: Compute normalized decision matrix from Step 1
Step 3: Load weightage vector with same value for all criteria
Step 4: Load impact vector with maximum as value for all criteria
Step 5: Compute TOPSIS matrix from normalized decision matrix, weightage vector and impact vector

The generated output using TOPSIS method ranks and identifies influencers in the given social media data set. Our proposed framework has flexibitlity to process social media data sets from Twitter, Facebook, Instagram etc. and it can be applied to various business domains such as healthcare, policitics, education etc.

4.1 Influence models

Influence user has the ability to convince others to accept a specific decision. This characteristic is based on on several factors, such as: popularity, social personality and influence abilities. Degree centrality, Eigen vector centrality and Page Rank are well known standard methods to understand the influence users. Various standard methods and our MCDM based TOPSIS method are summarized in Table 2.

4.1.1 MCDM based influence model

TOPSIS (technique for order preference by similarity to an ideal solution) [8, 16, 34] is one of the popular MCDM model is applied on the computed standard methods. TOPSIS method order the users based on given set of a multiple criteria measures computed. Extracted results from TOPSIS is combined with enterprise data to generate patterns interests.

4.2 Model formulation

Social media network data from Facebook, Twitter is shown in Fig. 2 and it is represented SMN = (G,P,L,C) where G is a graph descrbiing strcture of network as <V,E>, Set of postings of individuals in a network is denoted as P, L is the possible posting Likes, C is the possible of

Table 2 Influence models

Model	Explanation
Degree	Predict Influence users, who reaches other users directly
Eigenvector	Predict Influence users, who are well connected with other users
Page Rank	Predict Influence users who has high PageRank indicates the connection of high number of people connection
TOPSIS (MCDM)	Predict Influence users, close to the actual influencers from a possible set of influencers

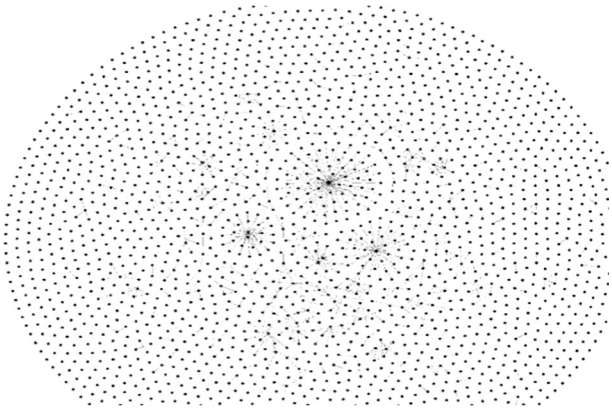


Fig. 2 Social network data representation in Gephi tool

comments posted by individual users, V is the vertices representing individual users and E is edges representing the relationship.

The degree C_j^{Degree} of node j is defined to be the degree of node:

$$C_j^{Degree} = \deg(\mathrm{j}) \tag{1}$$

The Eigenvector of a node a is defined,

$$\mathrm{x(a)} = \frac{1}{\lambda}\sum_{j=1}^{N} A_{a,b}\, x(b) \tag{2}$$

Where λ is a constant whereas adjacency matrix is described as A. The Eigenvector centrality understand not only the number of connections a vertex has but also the degree of the vertices which is connected to.

The Page Rank is defined as,

$$PR(u) = \frac{1-\propto}{N+\propto}\sum_{u\in M(u)} \frac{PR(u)}{L(u)} \tag{3}$$

Here u is node and L(u) is a function describes the group of individuals who liked the posting.

In TOPSIS, the relative closets of the alternative A_j with respect to A^+ is described as

$$R_j = d_j^- / \left(d_j^+ + d_j^- \right), \mathrm{j} = 1, .., \mathrm{m} \tag{4}$$

Since $d^-_j \geq 0$ and $d^+_j \geq 0$, then, clearly, $R_j \in [0, 1]$. Here A^+ is positive ideal solutions adjacency matrix, d^+_j and d^-_j are Euclidean distance of positive and negative ideal solutions.

5 Experiments and discussion

5.1 The dataset description

Around 20,000 tweets from #Demonetisation hashtag on December 9, 2016 between 11:29 and 18:53 IST from the Twitter public space was collected as social media



Fig. 3 Centrality measure calculation

dataset input to experiment the proposed framework. Using the algorithm proposed in section 4, an experiment was performed and compared the influence users identified. Twitter hashtagged #Demonetisation dataset is processed to compute various standard centrality results and MCMD based TOPSIS results for a SMA operation-Influence users identification using Gephi and R tool.

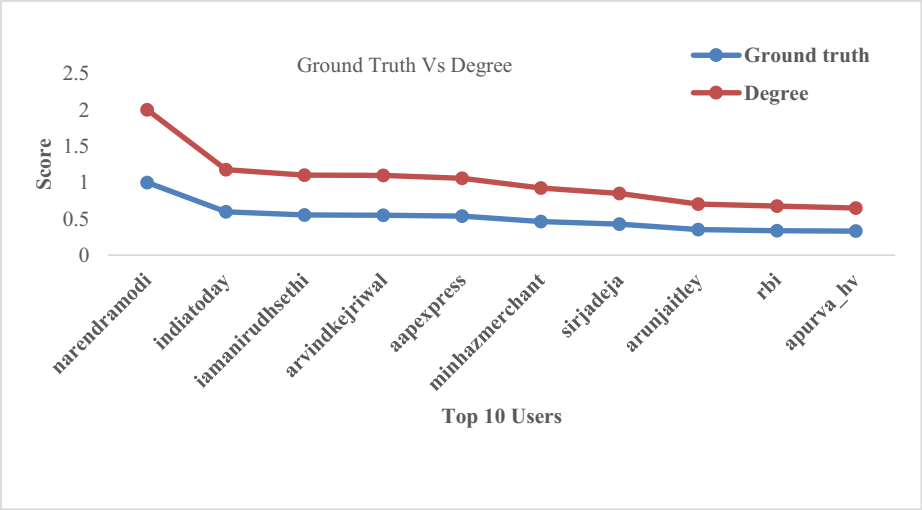


Fig. 4 Top 10 users comparison betweeen Ground Truth and Degree

5.2 Comparisons of influence users

Top 10 influence users by varous standard methods are listed in Table 3. The first three columns result from standard methods such as Degree, Eigen Vector Centrality, Page Rank whereas the last column result is based on MCDM based TOPSIS method. Many observations can be drawn by comparing and analyzing the influence users in Table 3.

1. Degree centrality and EigenVector centrality method determines the influencers (e.g. “narendramodi”, “arvindkejriwal”), who have large number of followers as shown in Figs. 4 and 5 respectively. Eventhough these influence users may not tweet often. One or two tweets of these influencers drive significant retweets or replies or mentions.

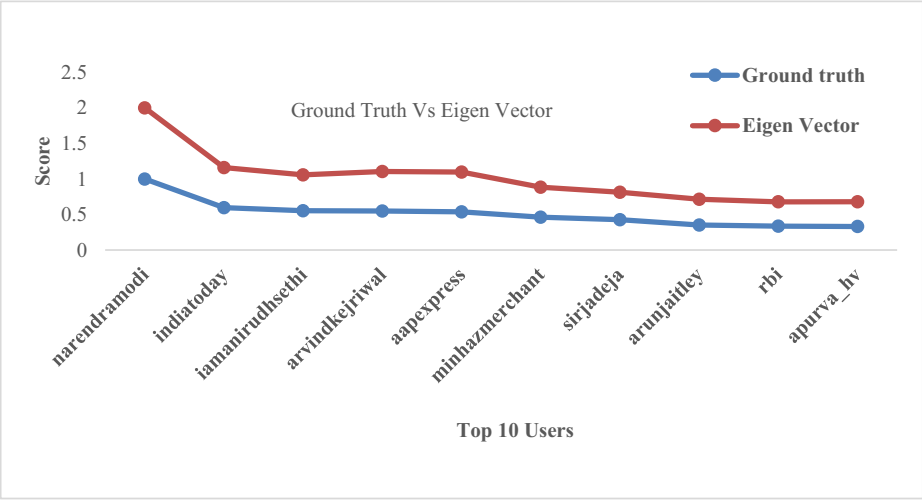


Fig. 5 Top 10 users comparison betweeen Ground Truth and Eigen Vector

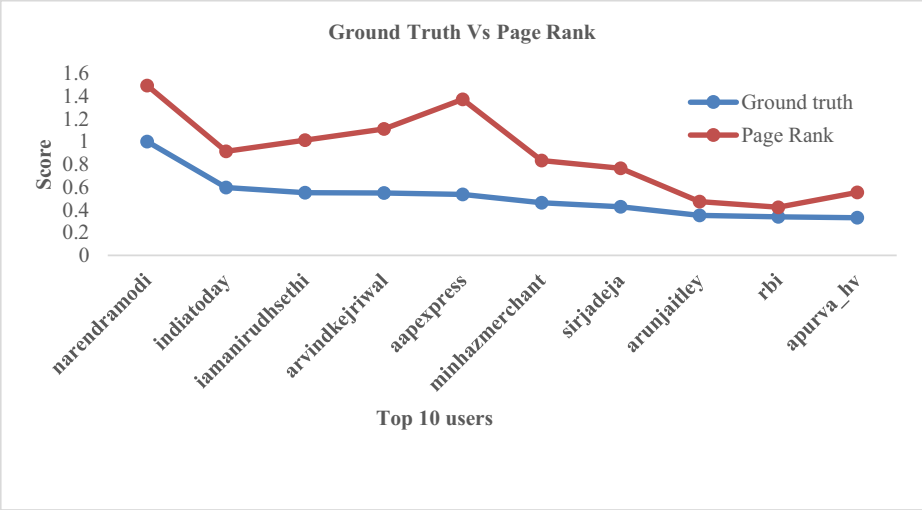


Fig. 6 Top 10 users comparison betweeen Ground Truth and Page Rank

2. PageRank determines top influencers not just from one-hop influence. PageRank is also determines influence users from whom they receive influence and subsequently spread in the network besides their direct followers as shown in Fig. 6. In the 3rd column of Table 1 lists the most influential users - famous celebrities with many followers and highly responsive influence users. The users like "DailyO" and "AAP Express" are popular news media frequently post reply and tweets or re-tweets with their followers.
3. The MCDM based TOPSIS method identifies the well-known influence users (e.g., "arvindkejriwal "narendramodi"), "along with popular news media (e.g. "indiatoday", ndtv, "AAP Express") as shown in Fig. 7. Our tweets were collated during the Demonetization's announcement by Mr. Narendra Modi, the Indian Prime

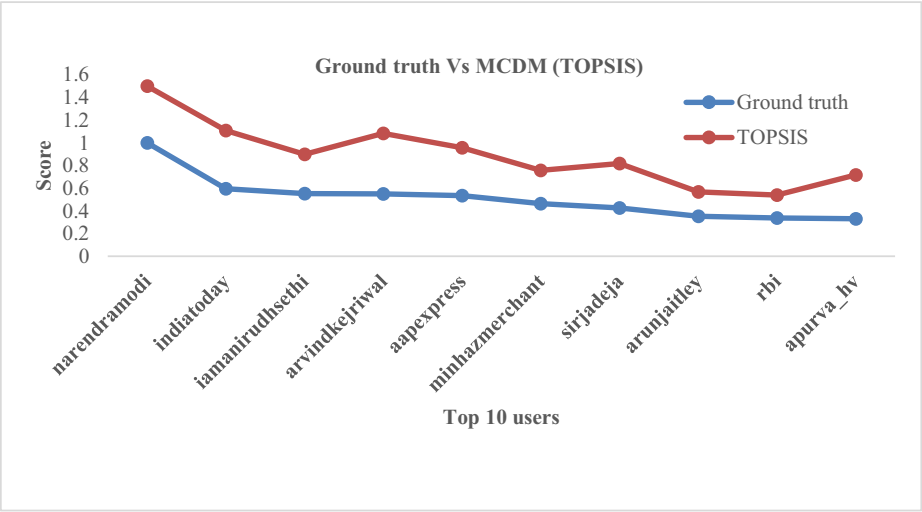


Fig. 7 Top 10 users comparison betweeen Ground Truth and MCDM (TOPSIS)

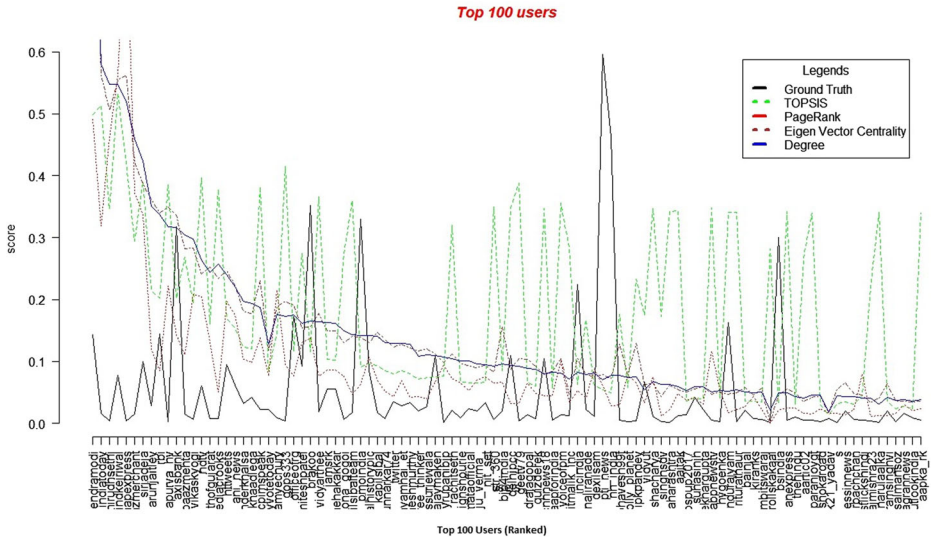


Fig. 8 Top 100 users Predicted Spreading Size with ground truth between TOPSIS Vs Standard methods

Minister. It is understood that MCDM based TOPSIS method has ability to identify influencers not only the celebrities who have many followers but also considering the user accounts in media, who post active replies, re-tweets and mentions.

5.3 Predicting spreading Size in TOPSIS- MCDM model

To prove our proposed TOPSIS (MCDM based) framework works better than existing standard methods, we calculated the Spreading Size of each user for a specific period of time

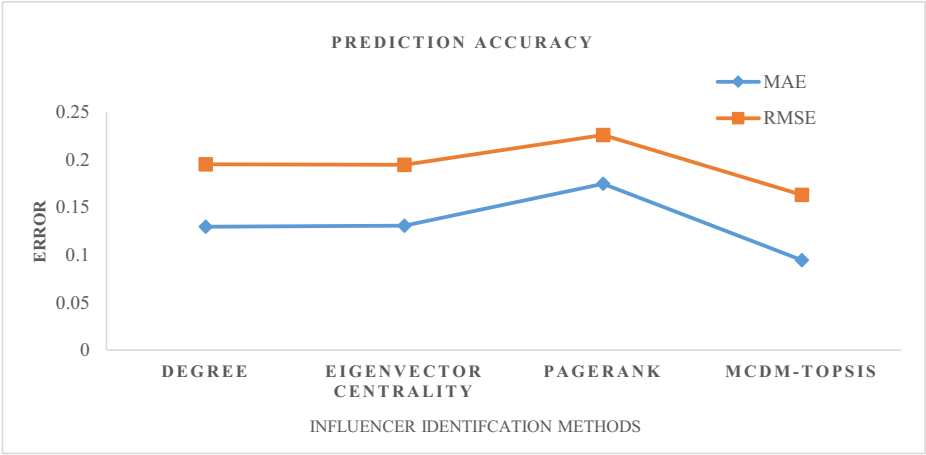


Fig. 9 Prediction Accuracy: Standard Methods Vs MCDM (TOPSIS)

to understand how many users are influenced over a particular topic. The next section explains how Spreading Size measure is used to predict the accurary of our proposed framework model.

5.3.1 The ground truth

The prediction accuracy of our framework and its feasibility is explained wit help of ground truth. Let a user m posts a topic ∝at t_1 time, soon after let n users post the same topic from t_1 to t_2 on the topic ∝, is denoted $SS_{m,\propto}^{t1\sim t2}$. For example, if a user A posts a topic on #Demonetisation hashtag in 2nd day, subseqently 10 more users are posting topic on #Demonetisation hashtag in 2nd day and 20 more users posting topics on #Demonetisation hashtag in 3rd day. The base line data of actual or ground truth Spreading Size (SS) for this post #Demonetisation hashtag to user A is 0 in 1st day, 10 in 2nd day and 20 in 3rd day. The average Spreading Size (SS) over all the topics for a user in a particular duration of time is computed as mentioned in eq. 5.

$$SS_m^{t1\sim t2} = \frac{\sum_{\propto} SS_{m,\propto}^{t1\sim t2}}{No.of\ \ topics(\propto)} \tag{5}$$

where m describes a user, t1 ~ t2 is a time window, and ∝ is a topic.

5.3.2 Prediction

The computed results from MCDM method is compared against standard methods on a specific SMA operation - Influence users identification. In Fig. 8, the ground truth Spreading Size is plotted with various methods for top 100 influence users. It shows how TOPSIS (MCDM based) prediction results along with standard methods fit with the ground truth or actual. It is observed that predicted results of TOPSIS (MCDM based) is very closely fit to the ground truth even though it is not the same with ground truth results, majority of the predicted curves matches closely align with ground truth. The other methods such as Degree, Eigen Vector centrality, Page Rank are nowhere close to the ground truth.

6 Performance accuracy

The performance of TOPSIS (MCDM based) results generated from the proposed framework is compared with standard methods results. Two well-known metrics- MAE (Mean Absolute

Table 3 The Top 10 influence users

Degree	Eigen Vector	Page Rank	TOPSIS (MCDM)
narendramodi	narendramodi	dailyo_	arvindkejriwal
indiatoday	indiatoday	prosaicview	indiatoday
arvindkejriwal	aapexpress	aapexpress	narendramodi
iamanidrudhsethi	arvindkejriwal	arvindkejriwal	dailyo_
aapexpress	iamanirudhsethi	narendramodi	prosaicview
minhazmerchant	minhazmerchant	iamanirudhsethi	aapexpress
sirjadeja	sirjadeja	minhazmerchant	gops333
arunjaitley	arunjaitley	sirjadeja	ndtv
rbi	apurva_hv	indiatoday	sirjadeja

Error), and RMSE (Root Mean Square Error) are applied to measure prediction accuracy as expressed in eq. 6 and 7.

Mean Absolute Error (MAE) measures the average magnitude of the errors in a set of predictions. It averages over the test sample of the absolute differences between prediction and actual observations where all individual differences have equal weight.

The MAE is given by,

$$\mathrm{MAE} = \sum_{i=0}^{n} |y_i - \hat{y}_i| = \frac{1}{n} \sum_{i=1}^{n} |e_i| \tag{6}$$

Where y_i *is actual and* $\hat{y}_i$ *is predicted*.

Root mean squared error (RMSE) is a quadratic scoring rule and it measures the average magnitude of the error. It is the square root of the average of squared differences between prediction and actual observations.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n} (y_i - \hat{y}_i)^2} \tag{7}$$

Where y_i *is actual and* $\hat{y}_i$ *is predicted*.

The average values of MAE and RMSE metrics for Degree, Eigen Vector, PageRank and MCDM – TOPSIS are listed in Table 4. It has been observed that TOPSIS (MCDM based) error values are least – 9% and 16% than the other standard methods as illustrated in Fig. 9.

7 Conclusion and future work

Our proposed Social Media Analytics based on Multi-Criteria Decision Making (MCDM) framework outperforms than standard methods for a SMA operation such as influence users identification. From the experiment results and its performance comparisons, it is found that the proposed model works better than the standard methods. The predicted Spreading Size (SS) of TOPSIS (MCDM based) achieves very promising performance as it match closely withe ground truth as well its performance accuracy is better than other standard methods. Our TOPSIS (MCDM based) framework showed least error accuracy and worked better than the standard methods.

For future work, the results extracted from SMA opertions such as recommending users and communities, Influence users, and top trending topics identification etc. can be further combined with enterprise data to refine the information insights for business decisions.

Table 4 The comparions over 10,000 tweets of top 100 users

Method	MAE		RMSE	
		in %		in %
Degree	0.1296941	13	0.195318	20
Eigen vector	0.1307366	13	0.194797	19
PageRank	0.1746722	17	0.225968	23
MCDM-TOPSIS	0.0945438	9	0.163059	16

References

1. Amiri M (2010) Project selection for oil-fields development by using the AHP and fuzzy TOPSIS methods. Expert Syst Appl 37(9):6218–6224
2. Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone's an Influencer: Quantifying Influence on Twitter. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, (WSDM '11), pp. 65–74
3. Behzadian M, Kazemzadeh R, Albadvi A, Aghdasi M (2010) PROMETHEE: A comprehensive literaturereview on methodologies and applications. Eur J Oper Res 200(1):198–215
4. Behzadian M, Otaghsara S, Yazdani M, Ignatius J (2012) A state-of-the-art survey of TOPSIS applications. Expert Syst Appl 39(17):13051–13069
5. Bentes A, Carneiro J, Silva J, Kimura H (2012) Multidimensional assessment of organizational performance: Integrating BSC and AHP. J Bus Res 65(12):1790–1799
6. P Brodka, "Key User Extraction Based on Telecommuniatoin Data (aka. Key Users in Social Network. How to find them?) 2013, arXiv:1302.1369
7. Cha M, Haddadi H, Benevenuto F, Gummadi PK (2010) Measuring User Influence in Twitter: The Million Follower Fallacy. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (ICWSM 2010), pp. 10–17
8. Chen SJ, Hwang CL (1992) Fuzzy Multiple Attribute Decision Making: Methods and Applications. Springer-Verlag, Berlin
9. Di Gangi PM, Wasko M (2009) Steal My Idea! Organizational Adoption of User Innovations From a User Innovation Community: A Case Study of Dell IdeaStorm. Decis Support Syst 48(1):303–3012
10. Domingos P, Richardson M (2001) Mining the network value of customers. ACM SIGKDD:57–66
11. Gandhi M, Muruganantham A (2015) Potential Influencers Identification using Multi-criteria Decision Making (MCDM) methods. Procedia Computer Science 57:1179–1188
12. Ghosh R, Lerman K (2010) Predicting Influential Users in Online Social Networks. In: SNA-KDD Proceedings of KSS workshop on Social Network Analysis. arXiv:1005.4882
13. Gobeck J, Hendler J (2006) Inferring binary trust relationships in web-based social networks ACM. TIOT: 497–529
14. Grul D, Guha R, Liben-Nowell D, Tomkins A (2004) Information diffusion through blogspace. ACM, In WWW:491–501
15. Han H, Trimi S (2018) A fuzzy TOPSIS method for performance evaluation of reverse logistics in social commerce platforms. Expert Syst Appl 103:133–145
16. Hwang CL, Yoon K (1981) Multiple Attribute Decision Making Methods and Applications. Springer, Berlin Heidelberg
17. Kaplan AM, Haenlein M (2010) Users of the World, Unite! The Challenges and Opportunities of Social Media. Business Horizons 53(1):61
18. Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network", In SIGKDD, pp. 137–146
19. Kiron D, Ferguson RB, Prentice PK (2013) From Value to Vision: Reimagining the Possible with Data Analytics. MIT Sloan Management Review: Spring Research Report, pp. 1–19
20. Kohavi R, Rothleder N, Simoudis E (2002) Emerging Trends in Business Analytics. Commun ACM 45(8): 45–48
21. Kumar A, Sah B, Singh AR, Deng Y, He X, Kumar P, Bansal RC (2017) A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. Renew Sust Energ Rev 69:596–609
22. Li L, Peng W, Kataria S, Sun T, Li T (2015) Recommending Users and communities in Social media. ACM Transactions on Knowledge Discovery from Data (TKDD) 10(17)
23. Li J, Peng W, Li T, Sun T, Li Q, Xu J (2014) Social network user influence sense-making and dynamics prediction. Expert Syst Appl 41(11):5115–5124
24. Li Q, Zhou T, Lu L, Chen D (2014) Identifying influential spreaders by weighted LeaderRank. PhysicaA: Statistical Mechanics and its Applications 404:47–55
25. Muruganantham A, Gandhi M (2016) Discovering and Ranking Influential Users in Social Media Networks Using Multi-Criteria Decision Making (MCDM) Methods. Indian J Sci Technol 9(32):1–11
26. Musial K, Kazienko P, Brodka P (2009) User Position Measures in Social Networks. In: Proceedings of the 3rd Workshop on Social Network Mining and Analysis. ACM, New York, No. 6.
27. Petz G, Karpowica M, Furchu H, Auinger A, Stritestky V, Holzinger A (2015) Computational approaches for mining user's opinions on the Web 2.0. Inf Process Manag:510–519
28. Richardson M, Domingos P (2002) Mining knowledge sharing sites for viral marking. SIGKDD:61–70
29. Riquelme F, Gonazalez P (2016) Measuring user influence on Twitter: A survey. Information Processing and Management 52(5):949–975

30. Shin H, Xu Z, Kim EY (2008) Discovering and Browsing of Power Users by Social Relationship Analysis in Large-Scale Online Communities. In Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology 1:105–111
31. X. Song, Y. Chi, K. Hino, B. Tseng (2007) Information flow modeling based on diffusion rate for prediction and ranking. In WWW, pp. 191–200
32. Song X, Tseng BL, Lin CY, Sun MT (2006) Personalized recommendation driven by information flow. SIGKDD:509–516
33. Spertus E, Sahami M, Buyukkokten O (2005) Evaluating Similarity Measures: A large-scale study in the Orkut social network. ACM SIGKDD:678–684
34. TOPSIS. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/TOPSIS#References.
35. Tran KT (2013) Introduction to Web Services with Java. Bookboon, London
36. Ul Mustafa R, Nawaz MS, Ullah Lali MI, Zia T, Mehmood W (2017) Predicting the cricket match outcome using crowd opinions on social Networks: A comparative study of machine learning methods. Malaysian Journal of Computer Science 30(1):63–76
37. Uzunkaya C, Ensari T, Kavurucu Y (2015) Hadoop Ecosystem and Its Analysis on Tweets. Procedia Soc Behav Sci 195:1890–1897
38. Vasuki V, Natarajan N, Lu Z, Savas B, Dhillon L (2011) Scalable Affiliation Recommendation using Auxiliary Networks. ACM Trans Intell Sys Technology
39. Velasquez M, Patrick T (2013) Hester, An Analysis of Multi-Criteria Decision Making Methods. International Journal of Operations Research 10(2):56–66
40. Vijayakumar A (2017) Automated risk identification using NLP in cloud based development environments. J Ambient Intell Humaniz Comput PP:1–13
41. Wang X, Triantaphyllou E (2008) Ranking irregularities when evaluating alternatives by using some ELECTRE methods. Omega 33(1):45–63

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Muruganantham A. is a Research scholar doing his Ph.D. with Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India and he received his M.E. in Computer Science and Engineering from National Institute of Technology (NIT), Tiruchirappalli, India in 2003 and his research focuses on Social Media data analytics.



G. Meera Gandhi is associated with Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India as a Professor in the Department of Computer Science and Engineering. Her research interests are Social Network Analysis, Data Mining, Applied Machine Learning and Databases, Deep Learning, Emotional Intelligence and Network Graph theories.

Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020)
IEEE Conference Record # 48766; IEEE Xplore ISBN: 978-1-7281-5371-1

Finger Recognition and Gesture based Virtual Keyboard

Chinnam Datta Sai Nikhil, Chukka Uma Someswara Rao, E.Brumancia, K.Indira, T.Anandhi, P.Ajitha
School of Computing

Sathyabama Institute of Science and
Technology Chennai, India

dattasai414@gmail.com, umasomesh.4455@gmail.com, easpinbrumancia@gmail.com, indirakaliannan@gmail.com,
anandhitamil2908@gmail.com, hannahgracelyne@gmail.com

***Abstract*—Hand motion acknowledgment is critical for human-PC connection. Right now, present a novel constant strategy for hand motion recognition. The proposed framework is vision based, which uses AI methods and contributions from a PC webcam. Vision based signal acknowledgment following and motion acknowledgment In our structure, the hand area is separated from the foundation with the foundation subtraction technique. At that point, fingers are portioned in order to identify and perceive the fingers. At long last, a standard classifier is applied to anticipate the names of hand motions. The examinations on the informational index of 1300 pictures show that our strategy performs well and is exceptionally productive. Besides, our technique shows preferred execution over a condition of-workmanship strategy on another informational collection of hand motions.**

Keywords—Hand Motion; Vision; Webca;,Finger recognition; gesture based;

I. INTRODUCTION

In this day and age, the PCs have become a significant part of life and are utilized in different fields notwithstanding, the frameworks and strategies that used to collaborate with PCs are obsolete and have different issues, which will talk about somewhat later right now. Consequently, an extremely new field attempting to defeat these issues has developed to be specific Human Computer Interactions (HCI). Despite the fact that, PCs have made various progression in the two fields of Software Hardware, Still the essential manner by which Humans collaborate with PCs continues as before, utilizing fundamental pointing gadget (mouse) and Keyboard or propelled Voice Recognition System, or possibly Natural Language handling in truly propelled cases to make this correspondence progressively human and simple for us.

Our proposed venture is the Hand motions acknowledgment framework to supplant the essential pointing gadgets utilized in PC to portray hand motions. Last, the affirmation of hand movements is rehearsed by evaluating the closeness of the component data.The input devices giving the main picture information fuses standard camera, sound framework camera, and ToF (time of flight) camera. The sound framework camera and ToF camera likewise give the significance information so it is definitely not hard to parcel the hand region from the establishment to the extent the significance map.

Right now, present a proficient and successful strategy for hand signal acknowledgment. The hand locale is frameworks to diminish the impediments that stay because of the heritage frameworks, for example, mouse and Touchpad. The proposed framework utilizes hand motion, for the most part no of fingers raised inside the locale of Interest to perform different tasks, for example, Play, Pause, look for forward, look for back word in video player. A static control board controls the adaptability of customer and limits the limit of the customer like a remote can be lost, dropped or broken while, the physical closeness of customer is required at sight of action and that is a confinement of the client. The proposed framework can be utilized to control different delicate boards like HMI frameworks, Robotics Systems, Telecommunication System, utilizing hand motions with assistance of programming by inside python utilizing pyautogui module to encourage cooperation inside various elements of PC through the Camera to catch video outlines.

To start with, the hand locale is recognized from the first pictures from the info gadgets. At that point, a few sorts of highlights are removed identified through the foundation subtraction strategy. At that point, the palm and fingers are part in order to perceive the fingers. After the fingers are perceived, the hand motion can be ordered through a straightforward principle classifier. The oddity of the proposed strategy is recorded as follows. (I) The principal oddity of the proposed technique is that the hand motion acknowledgment depends on the consequence of finger acknowledgment. Hence, the acknowledgment is cultivated by a straightforward and proficient principle classifier rather than the modern however confused classifiers, for example, SVM and CRF. (ii) Some past works need the clients to wear information glove to secure hand motion information. Be that as it may, the extraordinary sensors of information glove are costly and prevent its wide application, all things considered.

II. RELATED WORK

Most recent advancements in augmented reality (VR) innovation give amazing number of uses about wellbeing, amusement, restoration, instruction, and furthermore account. The main idea behind these

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:49:32 UTC from IEEE Xplore. Restrictions apply.

applications is to boost the Sagayam et al [10] propose some movement acknowledgment systems were concentrated on accessible movements with emphasis. Additionally, Hidden Markov model and variable classifier procedures over hand motion were given. Over the world, for conveying, gesture based communication is generally utilized by the hard of hearing network. For the hard of hearing individuals who experience the ill effects of hearing misfortune, it is extremely hard to speak with hearing individuals. In this manner, gesture based communication acknowledgment (SLR) was introduced to improve the correspondence.

Hassan et al. [11] explains Arabic SLR (ArSLR) was taken a shot at with Modified k- Nearest Neighbor and Hidden Markov Models (HMMs) methodology. In like manner, with Polhemus G4 development tracker and a camera, two new ArSLR datasets that have 40 Arabic sentences were gathered. Sign sentences assembled by development tracker were differentiated and sentences accumulated by sensor gloves on portrayal precision. As per the correlation, their grouping precision were fundamentally the same as. In Hasanuzzaman et al. [12], day by day exercises and medicine admission of elderly individuals was observed with radio recurrence ID (RFID) sensors and a camcorder at home. RFID perusers observed jugs. A camcorder checked the move of making medication with utilizing face discovery, foundation subtraction, action location and mouth identification. Another logical classifier was introduced to delineate vector to a character together with incline varieties identification in air-composing through a fingertip or deliver Mohammadi et al. [13]. Close by motion acknowledgment, neural systems are generally utilized for hand joints-based motion acknowledgment

In Ma et al. [14], to build up the precision close by signal acknowledgment from loud datasets, a settled interim unscented Kalman channel (UKF) together with long momentary memory (NIUKFLSTM) organize was introduced. Along these lines, the commotion in the consecutive hand skeletal information was rethought and the acknowledgment exactness was created. In the computer generated reality region, interoperability of virtual articles in augmented reality can be furnished by a Leap Motion with interfacing over an augmented simulation head- mounted presentation (VR HMD). Since Leap Motion can distinguish hand motions or precise finger positions. In any case, for an augmented experience application, it is uncomfortable to decide the specific situation of the real client's hands. Hence, mistakes develop during the cooperation of genuine client's hands and the virtual hands. In Park et al. [15], a strategy was introduced for organizing and planning the yield territory in VR HMD and detecting region in Leap Motion.

In Sivasangari et al .[15], The proposed architecture consists of a large convolutionary neural system (DCNN), where photographs and sketches1 are used to link exchange learning to enable the system to be comfortable with the connection between the two modalities.

In [16], The goal of this project is to develop an answer to a software system that identifies and classifies disease mechanically. The step such as loading an image, pre-processing, segmentation, extraction and classification involves the detection of disease. The photos of the leaves are used to identify the diseases of the plants. The goal of this project is to develop an answer to a software system that identifies and classifies disease mechanically. The step such as loading an image, pre-processing, segmentation, extraction and classification involves the detection of disease. The photos of the leaves are used to identify the diseases of the plants.

In Mary gladence et al [16] , A new model called (L, M, S, K) based on the Flexible and Accurate Motif Detector (FLAME) was introduced. FLAME is a versatile suffix-tree-based algorithm which can be used with a variety of meanings to find frequent patterns..

III. PROPOSED METHODOLOGY

Picture handling is a technique for conducting such procedures on a picture in order to get an improved picture or to retrieve any useful data from it. It is a kind of sign handling where information is a picture and yield may be image or attributes / highlights relevant to that picture.

Right now, against robbery gadget which would be sufficiently proficient to recognize burglary utilizing movement detecting camera utilizing AI and caution the proprietor with ready message alongside the caught picture of that occasion of movement.The Haar Cascade calculation is utilized to prepare information. The gadget will be a continuous framework alongside simple to utilize interface, which will be demonstrated helpful as far as security of individuals just as their important things/objects. Haar Cascade is an AI object discovery calculation used to distinguish protests in a picture or video.

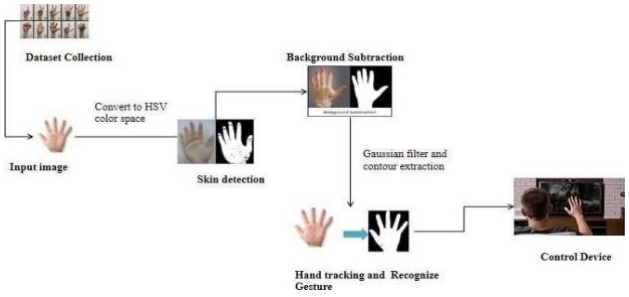


Fig. 1. Proposed architecture

A. Data Obtaining

The underlying move is to catch the picture from camera and to characterize an area of Interest in the edge, it is significant as the picture can contain a great deal of factors and these factors can bring about undesirable outcomes and the information that should be handled is diminished to a huge degree. To catch the picture a web-camera is utilized that ceaselessly catches outlines and is utilized to get the crude information for handling. The info picture have here is uint8. The Procured picture is RGB and must be handled before, for example, pre-prepared before the segments are isolated and affirmation is made.

B. Data preprocessing

The template Pre-preparing strategy can be finished 2- steps process: • Segmentation • Morphological sifting First Process is the Segmentation procedure. It is done to change over dark scale picture into the double picture so can have only two Are a of Interest in thepicture. That is, one will be hand and another is foundation. Calculation can be Otsu calculation' utilized for this procedure and dim scale picture are changed over into double picture having zone of enthusiasm as the hand and the foundation. There are two primary ways to deal with

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:49:32 UTC from IEEE Xplore. Restrictions apply.

division: 1. Pixel-based or nearby strategies having: - an) Edge discovery b) Boundary recognition. 2. Locale based methodology: Thresholding procedures is utilized in parceling the picture pixel histogram by utilizing a solitary limit system. Utilized otsu's thresholding procedure. Otsu's thresholding is utilized to consequently perform cluster-based thresholding.

The system generates an error in recognizing gesture this is termed as noise. It is used to avoid the noise and finish the contours of the hand motion in image

The strategy expect there are two classes of pixels following bi modular histogram and afterward it figures ideal edge isolating the two classes Morphological Filtering: After thresholding need to ensure that there will be no clamoris available in picture, so utilizing morphological sifting Techniques, These Techniques are partitioned in Also, there is a likelihood that framework created a mistake in perceiving motion this might be named as motion commotion. On the off chance that need perfect form identification of a motion, at that point previously mentioned blunders ought to be invalidated. A morphological isolating (separating) approach is utilized using gathering of expansion (augmentation) and disintegration (crumbling) to achieve a smooth, shut, and finish the shapes of a hand movement.

C. *Extraction of Features*

Pre-arranged or preprocessed picture is available to be used and various features of the resultant picture are evacuated. Following are the highlights that can be separated: I. Discovering Contours ii. Finding and adjusting curved frame iii. Numerical Operations Contours: It infers the bearing of hand for example whether or not the hand is on a level plane or vertically set. At first, attempt to discover the direction by length to width proportion with an assumption that if the hand is vertical ten length of the case jumping them will be more than the width of a similar box and, in the event that hand is on a level plane set, at that point width of bouncing box is bigger than width of the case restricting the hand will be more than that of the length of the case.Finding and remedying arched structures: A hand act is perceived by its own direction and by the way that what number of fingers are appeared. For getting the total of what number of fingers are appeared close by movements that need to process only region of the finger of the hand that have in past development by making sense of and breaking down the centroid.

IV. RESULTS AND DISCUSSION

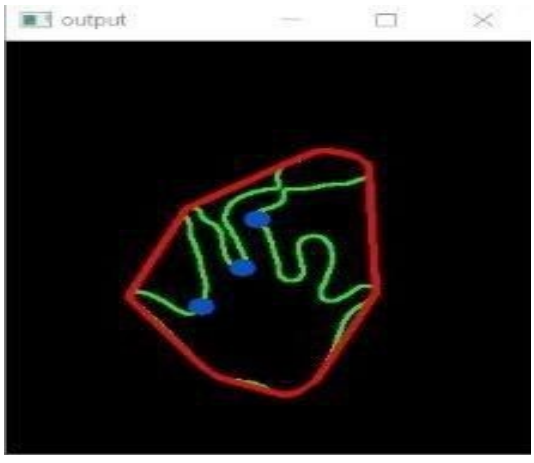


Fig. 2. Output image from Background Substraction

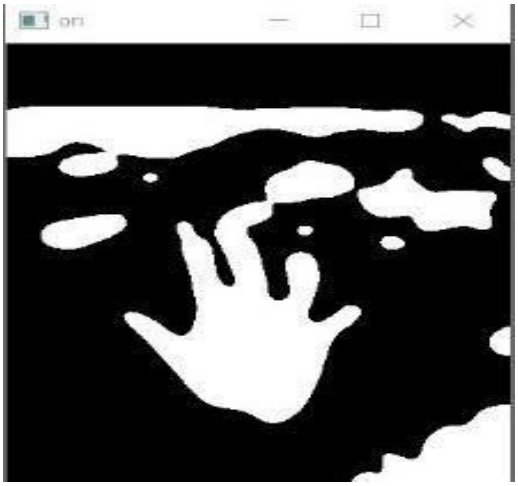


Fig. 3. Original image from Background Substraction



Fig. 4. Blur image from Background Substraction

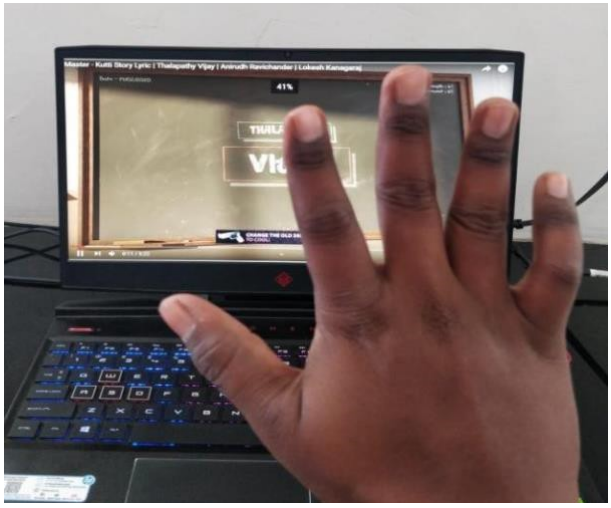


Fig. 5. Volume up Gesture

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:49:32 UTC from IEEE Xplore. Restrictions apply.



Fig. 6. Forward Gesture

The above figures shows that the how the gestures has bbeen used for control the video player. By using Background substraction method implementing the several filters to get the features of the original image. By showing single fingure to the webcam of the computer the action of the video player will be backward. BY showing Two fingers the action is volume decreased by showing three fingers the video will play forward and by showing five fingers the volume will be increased.

V. CONCLUSION

The vision based cursor control utilizing hand motion framework was created in the C++ language, utilizing the OpenCV library. The framework had the option to control the development of a Cursor by following the client's hand. Cursor capacities were performed by utilizing distinctive hand motions. The framework has the capability of being a practical trade for the PC mouse, anyway because of the imperatives experienced; it can't totally supplant the PC mouse. The significant requirement of the framework is that it must be worked in a sufficiently bright room. This is the principle motivation behind why the framework can't totally supplant the PC mouse, since it is normal for PCs to be utilized in open air situations with poor lighting condition. The precision of the hand motion acknowledgment could have been improved, if the Template Matching hand motion acknowledgment technique was utilized with an AI classified.Here, what did is, simply open our content document it will consequently dispatch a video player. Here picked VLC Media Player. At that point content stops execution for predefined time to stack the media player. After video document is being played then framework summons the instruments that required to run it for example OpenCV, Camera, pyautogui.

Presently, prepared to do simply kick back and control without utilizing any conventinal technique.

ACKNOWLEDGMENT

We are thankful to our management "Sathyabama Institute of Science and Technology" for giving a support to write this manuscript.

REFERENCES

[1] R. Zaman, K. Noor, A. Ibraheem, "Hand Gesture Recognition:A Literature Review", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012.

[2] L.Yun, Z. Lifeng, Z. Shujun, "A Hand Gesture Recognition Method Based on Multi- Feature Fusion and Template Matching", Procedia Engineering, Vol. 29, pp 1678-1684, 2012.

[3] S.S. Rautaray, A. Agrawal, "Real Time Gesture Recognition System for Interaction in Dynamic Environment," Procedia Technology, Vol. 4, pp 595- 599, 2012.

[4] Q. D. Smedt, H. Wannous, J.P.Vandeborre, "Heterogeneous hand gesture recognition using 3D dynamic skeletal data," Computer Vision and Image Understanding, Vol. 181, pp. 60-72, 2019.

[5] P. Caserman, A. Garcia-Agundez, R. Konrad, "Real- time body tracking in virtual reality using a Vive tracker", Virtual Reality (2019) 23: 155.

[6] K.M. Sagayam, D.J. Hemanth, "Hand posture and gesture recognition techniques for virtual reality applications: a survey", Virtual Reality, 2017, 21: 91.

[7] M. Hassan, K. Assaleh, T. Shanableh, " Multiple Proposals for Continuous Arabic Sign Language Recognition", Sens Imaging, 2019, 20: 4.

[8] F.M. Hasanuzzaman, X. Yang, Y. Tian, et al.," Monitoring activity of taking medicine by incorporating RFID and video analysis", Netw Model Anal Health Inform Bioinforma, 2013, 2: 61.

[9] S. Mohammadi, R. Maleki, ,"Real-time Kinect-based airwriting system with a novel analytical classifier", International Journal on Document Analysis and Recognition (IJDAR), 2, 2019.

[10] C. Ma , A. Wang , G. Chen, C. Xu, " Hand joints-based gesture recognition for noisy dataset using nested interval unscented Kalman filter with LSTM network", Vis. Comput. 2018, 34:1053–1063.

[11] C. Yang, J. Luo, Y. Pan, Z. Liu, and C.-Y. Su, "Personalized variable gain control with tremor attenuation for robot teleoperation," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 48, no. 10, pp. 1759- 1770, October 2018.

[12] V.I. Pavlovic, R. Sharma, T.S. Huang, "Visual interpretation of hand gestures for human- computer interaction: A review," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, issue 7, pp 677-695, 1997.

[13] F.S. Chen, C.M. Fu, C.L. Huang, "Hand gesture recognition using a real-time tracking method and hidden Markov models", Image and Vision Computing , vol. 21, issue 8, pp 745-758, 2003.

[14] M. Donmez, S. Dogan, E. Baran, "Sinyal İlkesinin Öğrenme Üzerindeki Etkileri: Bir Göz Hareketleri İzleme Çalışması," Mersin University Journal of the Faculty of Education, vol. 14 , issue 2, pp 700 – 713, 2018.

[15] Sivasangari, A., Poonguzhali, S., Rajkumar, I., Maheshwari ," Face photo recognition using sketch image for security system ", International Journal of Innovative Technology and Exploring Engineering , Vol. Volume-8, Issue- 9S2, July 2019 .

[16] Devaraj, A., Rathan, K., Jaahnavi, S., Indira, K, "Identification of Plant Disease using Image Processing Technique", 2019 International Conference on Communication and Signal Processing (ICCSP).

[17] L.Mary Gladence. Detection of Contiguous Pattern with Pattern Shift String Matcher. RESEARCH JOURNAL OF PHARMACEUTICAL BIOLOGICAL AND CHEMICAL SCIENCES. 2017 Mar 1;8(2):2449-56.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:49:32 UTC from IEEE Xplore. Restrictions apply.

Feature Selection Using Multiple Kernel Learning Methods



K. Siva Sandeep Reddy, K. Gunavardhan Reddy, A. Pravin, G. Nagarajan, and T. Prem Jacob

1 Introduction

Kernal-based bunching calculations, for example, *k*-mean part, can catch the characteristic nonlinear structure in numerous genuine world datasets and consequently by and large perform preferred in bunch execution over direct parcel techniques [1]. In genuine applications, we can assemble a few centers and apply a few principle capacities. Be that as it may gathering execution to a great extent relies upon the decision of grains [2]. Sadly, it is as yet a test to decide a reasonable among a wide scope of potential centers for information gave and undertakings ahead of time. It is increasingly troublesome, particularly in solo learning exercises, for example, gathering, because of the nonattendance of labels. To conquer this trouble, numerous solo different portion learning strategies are accessible proposed [2, 3], which plan to become familiar with a center of agreement of a gathering of client characterized centers. Customary multi-center learning strategies without supervision gain proficiency with an accord center by directly joining a set of applicant beans. For instance, [4] gave a strategy for otherworldly gathering of numerous perspectives in a direct manner joined phantom embeddings to get the last bunching. The technique proposed a restricted various parts *k*-implies strategy for malignant growth

K. Siva Sandeep Reddy (✉) · K. Gunavardhan Reddy · A. Pravin · G. Nagarajan · T. Prem Jacob
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India
e-mail: Kssr143.143@gmail.com

K. Gunavardhan Reddy
e-mail: gunavardhan8285@gmail.com

A. Pravin
e-mail: Pravin_ane@rediffmail.com

G. Nagarajan
e-mail: nagarajanme@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2021

K. S. Sherpa et al. (eds.), *Advances in Smart Grid and Renewable Energy*,
Lecture Notes in Electrical Engineering 691,
https://doi.org/10.1007/978-981-15-7511-2_56

science applications [5]. A different piece learning technique is learning the ideal neighborhood for the bunching [3].

The above strategies focus on consolidating all up-and-comer bits, while overlook the vigor of the techniques. In any case, in certifiable applications, the bits are regularly tainted by clamors. For instance, since the first information may contain clamors and exceptions, the piece built with these information will likewise be defiled. What is more, despite the fact that the information is spotless, ill-advised piece capacities may likewise present clamors.

2 Related Works

To ease the exertion of commotions, as of late, some strong different piece learning strategies [6, 7] are proposed. These techniques center around the commotions brought about by the defiled occurrences, though cannot catch the commotions initiated by piece works well. Note that, some different part learning techniques dole out bigger load on the more suitable parts can catch the commotion incited by bit capacities to a few degree. Be that as it may, in these techniques, the weight is forced on all components of the piece lattice, and it is a piece excessively harsh. For model, some wrong pieces may have a low weight in these techniques, which implies all components including the helpful parts in the piece lattice will have a similar low weight and won't be useful to the kernel-based learning. To deal with commotions all the more extensively, right now, we propose a novel local and global de-noising multiple bit learning technique. We see that the piece grid may contain two sorts of clamors: one is brought about by debased occurrences, and the other is brought about by unseemly portion capacities. Primary proposed by Yohai (1988), MM-estimators have become increasingly popular and perhaps now in the most commonly employed robust regression technique. They combine a heavy breakdown point (50%) with good efficiency (approximately 95%). The "MM" in the name refers to the fact that more than one *M*-estimation procedure is used to calculate the final estimates. It has both the high breakdown property and a higher statistical efficiency than S estimation.

3 Proposed Work

Multiple portion learning has been effectively examined [3, 5, 8]. In view of the accessibility of class marks, different piece learning can be classified into two classes: directed calculations and unaided strategies. In supervised learning, names of occasions are accessible, which can be useful to take in an agreement portion from different parts. For instance, Liu et al. coordinated range data into numerous part learning to improve the bit learning execution; stretched out extraordinary learning machine to various portion picking up prompting a numerous bit outrageous learning

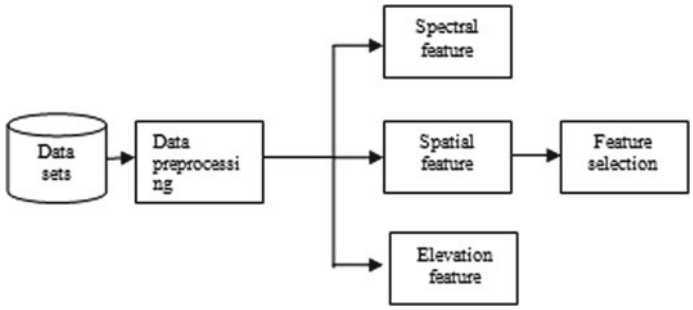


Fig. 1 Architectural view

machine. Albeit administered various portion learning has been broadly contemplated, solo calculation is all the more testing because of the nonappearance of class names. In unaided learning, a few techniques broaden single piece-based bunching strategy into various bit setting. For instance, proposed a different part fluffy grouping techniques by presenting a network initiated regularization and gave a self-weighted different part learning calculation and applied it to a diagram-based bunching technique. Since part k-implies is a well-known bunching strategy, numerous unaided different piece learning techniques have P. Zhou et al.: Unsupervised robust multiple kernel Learning by means of extracting local and global noises been created in its system. They likewise proposed two co-regularization-based methodologies for multi-see ghostly grouping by upholding the bunching theories on various perspectives to concur with one another collected pieces with various loads into a brought together one for ghastly grouping. By falling back on the otherworldly technique, additionally proposed to unravel a different part k-implies related with two-layer loads [5]. Like [5, 6] first introduced a various bit k-implies strategy which considered the portion connections, and afterward relying upon the outcomes in [1], they move the part *k*-intends to phantom strategies furthermore, get the grouping results from eigenvalue decay.

Figure 1 gives the architectural view of the system. The datasets that are collected are grouped, and it will be given for the next stage. The next process will be the data preprocessing. The output will be the extraction of the special feature and the elevation feature, and finally, the feature selection is performed.

4 Results and Discussion

We provide in the reference color scene in Fig. 2 and after the computation we get noisy low-resolution depth map and after that we get reconstructed high-resolution depth map.

We provide the black and white image in Fig. 3 containing some noise and specify the user required colors and after the execution we obtain the desired outcome.

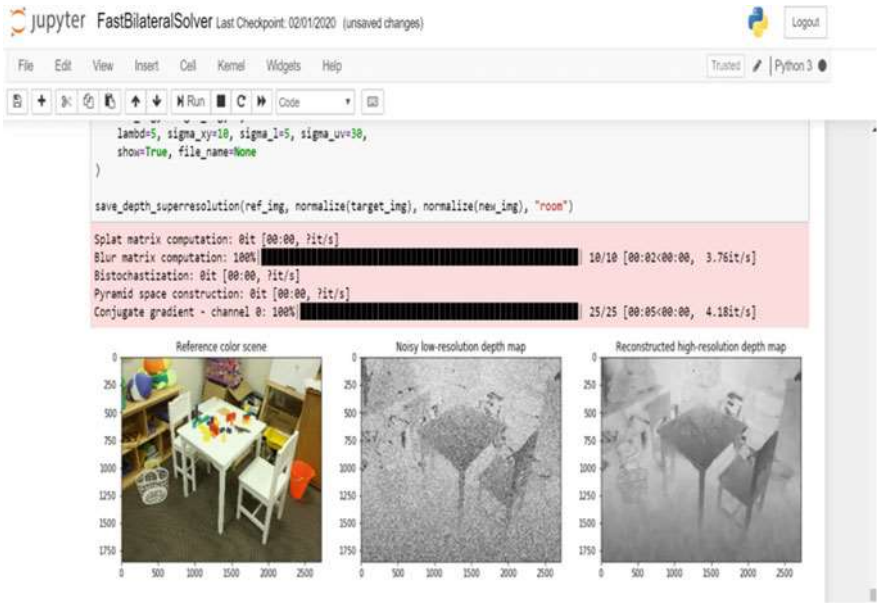


Fig. 2 Reference color scene

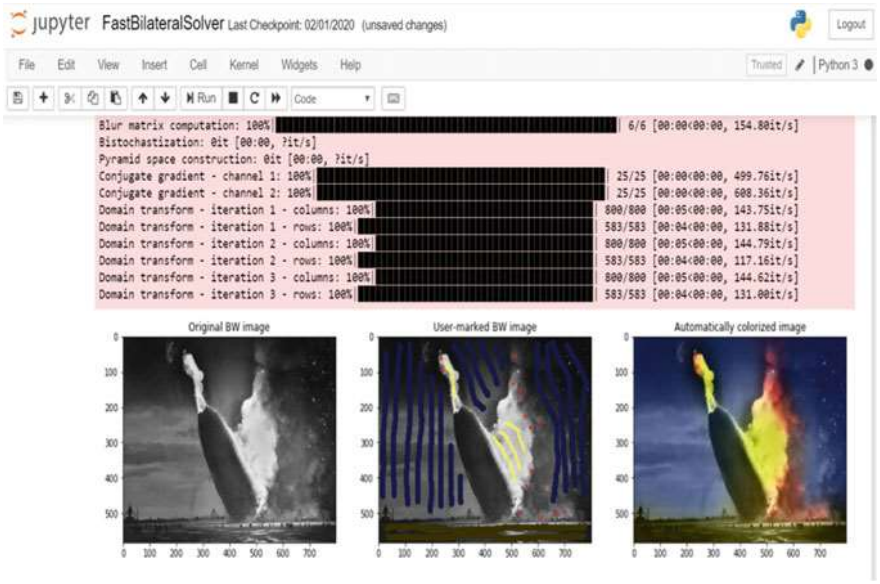


Fig. 3 Original BW image to automatically image

5 Conclusion

Contrasted and the isotropic dispersion model and the absolute variety model, the new model inferred right now dependent on PDEs. It conquers the staircase impact, which is inalienable in the idea of the picture denoising by the isotropic dissemination in the level area, while secures the limit highlights; it defeats the square impact, which is innate in the idea of the picture denoising by the all-out variety model, while holds the neighborhood subtleties and qualities. From the presentation of the reenactments, the new model has more denoise capacity regarding PSNR contrasted with the isotropic dispersion model and the absolute variety model. Note that numerical estimations do not consider other proficient counts and progressively appropriate weight capacities. On the off chance that different techniques for comprehending the halfway differential conditions are applied to the new model, the intermingling rate might be improved. Furthermore, because of the absence of the earlier information on the picture and the investigation of the general structure of the picture, the exhibition of denoising of the new model should be improved, and these are inspired us for future research.

References

1. K. Dabov, A. Foi, V. Katkovnik, Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
2. Y.J. Li, J. Zhang, J. Wang, Improved BM3D denoising method. IET Image Process. **11**(12), 1197–1204 (2017)
3. L. Du et al., Robust multiple kernel *k*-means using 2,1-norm, in *Proceedings of the 4th International Joint Conference on Artificial Intelligence*, Jun 2015, pp. 3476–3482
4. X. Liu, Y. Dou, J. Yin, L. Wang, E. Zhu, Multiple kernel *k*-means clustering with matrix-induced regularization, in *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, Feb 2016, pp. 1888–1894
5. X. Liu, M. Li, L. Wang, Y. Dou, J. Yin, E. Zhu, Multiple kernel *k*-means with incomplete kernels, in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, Feb. 2017, pp. 2259–2265
6. M. Gönen, A.A. Margolin, Localized data fusion for kernel *k*-means clustering with application to cancer biology. Proc. NIPS **2**, 1305–1313 (2014)
7. V.R. Sarma Dhulipala, P. Devadas, P.H.S. Tejo Murthy, Mobile phone sensing mechanism for stress relaxation using sensor networks: a survey. Wirel. Pers. Commun. **86**, 1013–1022 (2016). https://doi.org/10.1007/s11277-015-2969-y
8. X.J. Mao, C. Shen, Y.B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections (2016). arXiv preprint arXiv:1603.09056
9. J. Xie, L. Xu, E. Chen, Image denoising and inpainting with deep neural networks, in *Advances in Neural Information Processing Systems* (Nevada, USA, 2012), pp. 341–349
10. F. Zhang, N. Cai, J. Wu, G. Cen, H. Wang, X. Chen, Image denoising method based on a deep convolution neural network. IET Image Processing, **12**(4), 485–493 (2018)

Detection of Ransomware Based on Recurrent Neural Network (RNN)



Saketh Kanumuri, Vinay Teja Kantipudi, A. Viji Amutha Mary, and Mercy Paul Selvan

1 Introduction

With the expanded utilization of PC frameworks and the development of the Internet, cybercrime has additionally expanded [1]. In the created world, processing and Web use is imbued in the public eye, the economy and foundation. This perplexing blend has made a rewarding situation of potential unfortunate casualties, which can be gone after by cybercriminals [2]. Digital assaults are expanding day by day, regularly powered by advancing innovation [3, 4]. Ransomware has developed as a significant danger to people and organizations the same. Given the pervasive idea of figuring, ransomware is regularly profoundly computerized, consequently empowering lawbreakers to focus on a wide scope of exploited people like ventures, administrative organizations and people [5].

Ransomware is getting progressively hazardous, causing critical harm around the world. Ransomware—a compound expression of payoff and programming—is utilized to encode and incapacitate a client's important information, expecting instalment to recuperate it. The high identification exactness of grouping learning models is for the most part fuelled by the capacity of repetitive neural systems, for example, a Long Short-Term Memory (LSTM) [6], to tie between relatedness

S. Kanumuri · V. T. Kantipudi · A. V. A. Mary (✉) · M. P. Selvan
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India
e-mail: vijiamumar@gmail.com

S. Kanumuri
e-mail: saketh.kanumuri@gmail.com

V. T. Kantipudi
e-mail: vinaytejakantipudi@gmail.com

M. P. Selvan
e-mail: mercypaulselvan.cse@sathyabama.ac.in

© Springer Nature Singapore Pte Ltd. 2021

K. S. Sherpa et al. (eds.), *Advances in Smart Grid and Renewable Energy*,
Lecture Notes in Electrical Engineering 691,
https://doi.org/10.1007/978-981-15-7511-2_57

between occasions happening in a specific consecutive request. While such request can be caught correspondingly in ransomware successions, we accept that these arrangements display certain extra properties.

The ransomware scrambles the documents surprisingly fast once it begins activity making the records out of reach [7, 8]. A strategy that investigations the executables even before they start is to be created [9]. The proposed work does an itemized investigation of API calls as it tends to be utilized to perform malignant exercises. Since API calls are the essential square of a program, its examination can be utilized for seeing how a record acts. The work done right now is as follows:

- Dealing with class awkwardness as the quantity of ransomware tests are a lot higher than kind examples
- Applying different AI classifiers to distinguish the best model for identifying ransomware
- Analysis of a colossal corpus of ransomware and generous documents to recognize the most distinctive API calls among ransomware and amiable records.

2 Relatedwork

The present assemblage of writing contains some work concentrating on the investigation and discovery of ransomware assaults dependent on arrange traffic. Cabaj et al. [10] dissected Crypto ransomware powerfully in a devoted domain utilizing a honeypot and programmed run-time malware systematic framework. They recognized a portion of the system exercises of Crypto and introduced functional outcomes, i.e. distinguishing some intermediary servers, the conventions utilized and the hard-coded addresses of certain servers. It does not remove any nonexclusive social component that can be applied to the recognition of different sorts of Crypto ransomware. Ahmadian et al. [11] proposed the Connection Monitor and Connection Breaker (CM&CB) system to recognize 'high survivable ransomwares' which utilize Domain Generation Algorithm (DGA). The proposed strategy examines the mentioned area names against a recently prepared English language Markov chain model. This work is assembled exclusively upon a solitary system highlight. Further, the viability of utilizing a Markov chain model against short messages, i.e. an area name, is faulty and may cause high bogus positive rates. This technique requires the structure and support of Markov chain models for some non-English dialects, which would be a costly procedure. Ramya et al. proposed a technique for restricting mischievous users in anonymous networks [12].

Kharraz et al. [13] examined ransomware having a place with different classifications like work area locking and those utilizing encryption procedures for locking the client documents. They arrived at the determination that there are various ransomware which show comparative examples and they can be separated from ordinary amiable processes. They recommended procedures dependent on observing of record framework exercises; however, these techniques were not assessed by the creators Kharraz et al. has recommended strategies which check for the similitude

of screen captures that are taken before executing a ransomware and subsequent to executing a ransomware which has a place with the storage family. Their techniques do not proactively recognize ransomware when it enters the framework and starts its activity. An elective reinforcement office is likewise excluded from their investigation.

2.1 Network Traffic Analysis

$$\begin{aligned} \mathrm{Rt} &= \mathrm{MATRIX}(\mathrm{St}) \\ \mathrm{Mt} &= \mathrm{DENSE}(\mathrm{Wd} * \mathrm{Rt}) \end{aligned} \tag{1}$$

$$\begin{aligned} \alpha t &= \mathrm{SoftMax}(\omega \mathrm{TMt}) \\ \mathrm{rt} &= \mathrm{Rt}\,\alpha\,\mathrm{Tt}. \end{aligned} \tag{2}$$

3 Proposedsystem

The approach for the investigation and arrangement of the different kinds of record in our work is talked about underneath. When an executable is gotten by the framework, our model ought to have the option to mark it as (a) Ransomware, i.e. malevolent programming that scrambles the client records. Benign applications are not pernicious. The fundamental plan to accomplish this objective is by a static investigation of the executable and utilizing API calls utilized as the most separating highlight to recognize ransomware and amiable documents. The various advances done to land at this are (1) Feature extraction and investigation: We remove out the API calls utilized in the different kinds of executables, and a point by point examination is done to recognize the most separating sets of highlights to be utilized for arrangement. (2) Dealing with class unevenness: Class awkwardness happens when there are extensively a larger number of instances of one objective class than another. We have an imbalanced dataset as the ransomware class is the dominant part. This is a significant issue looked by numerous malware location frameworks [13, 14]. Class lop-sidedness may influence the working of the classifier, and there is an inclination towards the greater part class. We have managed this issue by applying Smote method by chawla et al. [3]. (3) Classification: The information removed from the component extractor is sent to an essential classifier like choice tree or troupe of classifiers like an irregular woodland for preparing. At the point when another st of test information is given, the classifier yields the forecast whether the given record is ransomware or not. The proposed model is appeared in Fig. 1. The executable document which might be a ransomware or an amiable record initially experiences a

Springer Link

Search Log in



Artificial Intelligence Techniques for Advanced Computing Applications
pp 249–258

Feature Selection Techniques for Disease Diagnosis System: A Survey

G. Saranya & A. Pravin

Conference paper | First Online: 24 July 2020

482 Accesses | **2** Citations

Part of the Lecture Notes in Networks and Systems book series (LNNS,volume 130)

Abstract

Reducing dimensionality as a preprocessing phase towards artificial intelligence effectively eliminates unnecessary and repetitive information, raises the efficiency of learning accuracy, and improving output understandability. Nonetheless, the rapid surge in data dimension presents a real challenge to several current collections of features and extraction methods as regards efficiency and effectiveness. Many researchers build and use plenty of feature selection algorithms. But an emerging field of machine learning is still to be based on data mining and method of analysis for information processing. Due to the recent increase in data variation and speed, many feature selection

algorithms face serious efficiency and performance problems. Different kinds of feature selection algorithms are available in research such as algorithms based on filters, algorithms based on wrappers, and algorithms based on hybrids. Additionally, a literature study analyzes some of the existing popular feature selection algorithms and also addresses the strong points and difficulties of those algorithms.

Keywords

Feature selection **Machine learning**

Dimensionality reduction

Feature optimization

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-5329-5_24
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 117.69
Price includes VAT (India)

- ISBN: 978-981-15-5329-5
- Instant PDF download
- Readable on all devices

- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 139.99

Price excludes VAT (India)

- ISBN: 978-981-15-5331-8
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 199.99

Price excludes VAT (India)

- ISBN: 978-981-15-5328-8
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. Chumerin N, Hulle MMV (2006) Comparison of two feature extraction methods based on maximization of mutual information. In: Proceedings of the 16th IEEE signal processing society workshop on machine learning for signal processing, pp 343–348

2. Motoda H, Liu H (2002) Feature selection, extraction and construction. In: Towards the foundation of data mining workshop, sixth Pacific-Asia conference on knowledge discovery and data mining (PAKDD2002), Taipei, Taiwan, pp 67–72

3. Mhamdi H, Mhamdi F (2014) Feature selection methods on biological knowledge discovery and data mining: a survey. In: 2014 25th international workshop on database and expert systems applications, pp 46–50. IEEE

4. Janecek AGK, Gansterer GF et al (2008) On the relationship between feature selection and classification accuracy. In: Proceeding of new challenges for feature selection, pp 40–105

5. Khalid S, Khalil T, Nasreen S (2014) A survey of feature selection and feature extraction techniques in machine learning. In: 2014 science and information conference, pp 372–378. IEEE

6. Xue B, Zhang M, Browne WN, Yao X (2015) A survey on evolutionary computation approaches to feature selection. IEEE Trans Evol Comput 20(4):606–626

7. Liu X, Shang L (2013) A fast wrapper feature subset selection method based on binary

particle swarm optimization. In: 2013 IEEE congress on evolutionary computation

8. Asir Antony Gnana Singh D, Appavu alias Balamurugan S, Jebamalar Leavline E (2016) Literature review on feature selection methods for high-dimensional data. Int J Comput Appl (0975–8887) February 2016

9. Veerabhadrappa L, Rangarajan (2010) Bi-level dimensionality reduction methods using feature selection and feature extraction. Int J Comput Appl 4(2):33–38

10. Pereira LAM, Rodrigues D, Almeida TNS, Ramos CO, Souza AN, Yang XS, Papa JP (2014) A binary cuckoo search for feature selection. Springer International Publishing Switzerland

11. Nakamura RYM, Pereira LAM, Costa KA, Rodrigues D, Papa JP, Yang X-S (2011) BBA: a binary bat algorithm for feature selection. In: Tangri N, Stevens LA, Griffith J, Tighiouart H, Djurdjev O, Naimark D, Levin A, Levey AS (eds) 2012 XXV SIBGRAPI conference on graphics, patterns and images. A predictive model for progression of chronic kidney disease to kidney failure. JAMA 305(15), 1553–1559 (2011)

12. Spolaor N, Carolina Lorena A, Diana Lee H (2010) Use of multi objective genetic algorithms in feature selection. In: 2010 eleventh Brazilian symposium on neural networks

13. James Mathai K, Agnihotri K (2017) Optimization techniques for feature selection in classification. 2017 IJEDR | Volume 5, Issue 3 | ISSN: 2321-9939

14. Soliz P et al (2008) Independent component analysis for vision-inspired classification of retinal images with age-related macular degeneration. In: Proceeding of IEEE international conference on image processing SSIAI, pp 65–68

15. Ladla L, Deepa T (2011) Feature selection methods and algorithms. Int J Comput Sci Eng (IJCSE) 3(5):1787–1797

16. Zheng Y et al (2013) An automated drusen detection system for classifying age-related macular degeneration with color fundus photographs. In: IEEE 10th international symposium on biomedical imaging, pp 1440–1443

17. Tan et al (2009) A hybrid evolutionary algorithm for attribute selection in data mining. J Expert Syst Appl 36:8616–8630

18. Yan H et al (2008) Selecting critical clinical features for heart diseases diagnosis with a real-coded genetic algorithm. J Appl Soft Comput 8:1105–1111

19. Bhatia N et al (2012) An analysis of heart disease prediction using different data mining techniques. Int J Eng Res Technol 1(8)

20. Quinlan JR (1996) Improved use of continuous attributes in C4.5. J Artif Intell Res 4:77–90

21. Hamilton HJ, Shan N, Cercone N (1996) RIAC: a rule induction algorithm based on approximate classification. Computer Science Department, University of Regina

22. Ster B, Dobnikar A (1996) Neural networks in medical diagnosis: comparison with other methods. In: Proceedings of the international conference on engineering applications of neural networks (EANN '96), pp 427–430

23. Peña-Reyes CA, Sipper M (1999) A fuzzy-genetic approach to breast cancer diagnosis. Artif Intell Med 17(2):131–155

24. Akay MF (2009) Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst Appl 36(2):3240–3247

25. Huang C-L, Liao H-C, Chen M-C (2008) Prediction model building and feature selection with support vector machines in breast cancer diagnosis. Expert Syst Appl 34(1):578–587

26. Chen H-L, Yang B, Liu J, Liu D-Y (2011) A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. Expert Syst Appl 38(7):9014–9022

27. Rathore N, Agarwal S (2014) Predicting the survivability of breast cancer patients using ensemble approach. In: Proceedings of the international conference on issues and challenges in intelligent computing techniques (ICICT ’14), February 2014. IEEE, pp 459–464

28. Polat K, Güneş S (2007) Breast cancer diagnosis using least square support vector machine. Digit Signal Proc 17(4):694–701

29. Krishnaveni N, Radha V (2019) Feature selection algorithms for data mining classification: a survey. Indian J Sci Technol 12(6). https://doi.org/10.17485/ijst/2018/v12i6/139581, February 2019

30. Chandrashekar G, Sahin F (2014) A survey on feature selection methods. Comput Electr Eng 40(1):16–28

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

G. Saranya & A. Pravin

Corresponding author

Correspondence to G. Saranya .

Editor information

Editors and Affiliations

Karunya Institute of Technology and Sciences, Coimbatore, India

Dr. D. Jude Hemanth

SRM Institute of Science and Technology, Chennai, India

Dr. G. Vadivu

SRM Institute of Science and Technology, Chennai, India

Dr. M. Sangeetha

Aurel Vlaicu University of Arad, Arad, Romania

Prof. Dr. Valentina Emilia Balas

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Saranya, G., Pravin, A. (2021). Feature Selection Techniques for Disease Diagnosis System: A Survey. In: Hemanth, D., Vadivu, G., Sangeetha, M., Balas, V. (eds) Artificial Intelligence Techniques for Advanced Computing Applications. Lecture Notes in Networks and Systems, vol 130. Springer, Singapore. https://doi.org/10.1007/978-981-15-5329-5_24

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-5329-5_24

Published	Publisher Name	Print ISBN
24 July 2020	Springer, Singapore	978-981-15-5328-8

Online ISBN	eBook Packages
978-981-15-5329-5	Intelligent Technologies and Robotics

Intelligent Technologies and Robotics (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Springer Link

Search Log in



Advances in Data Science and Management pp 67–79

Feature Selection and Clustering of Documents Using Random Feature Set Generation Technique

A. Christy & G. Meera Gandhi

Conference paper | First Online: 14 January 2020

652 Accesses | **1** Citations

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 37)

Abstract

In this digital world, information exists in abundance everywhere. Document clustering and instances involving pattern mining requires a subset of relevant features for analysis. Feature sets, which constitute one or more words, from the content of a document play a vital role in document clustering. A filter-based method for feature selection algorithm called random feature set generation (RFG) is proposed in this paper for document clustering. In feature selection, the selected features are checked for its quality using a quality metric and the best quality terms are used as the basis for

document clustering ignoring the worst terms as well as frequently occurring common terms from the corpus. Exhaustive experimentation for identification of feature sets is impossible due to its increased demand for computational effort. The advantage of RFG approach lies in selecting “good” quality terms for document clustering. The feature sets thus obtained are used for document clustering using K-means and X-means clustering algorithms. Experimental results have shown “good” quality terms filtered using random feature set generation (RFG) do not rely on monotonicity assumptions, and it shows positive correlation than the feature sets obtained using sequential forward feature set generation.

Keywords

Feature sets **Clustering**

Sequential forward **Random feature set**

K-means **X-means**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-0978-0_6
- Chapter length: 13 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 128.39

Price includes VAT (India)

- ISBN: 978-981-15-0978-0
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 159.99

Price excludes VAT (India)

- ISBN: 978-981-15-0980-3
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 219.99

Price excludes VAT (India)

- ISBN: 978-981-15-0977-3
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. H. Almuallim, T.G. Dietterich, Learning boolean concepts in the presence of many irrelevant features. Artif. Intell. **69**(1–2), 279–305 (1994)

2. M. Dash, H. Liu, Consistency-based search in feature selection. Artif. Intell. **151**(1–2), 155–176 (2003)

3. M.A. Hall, Correlation-based feature selection for discrete and numeric class machine learning, in *Proceedings of 17th International Conference on Machine Learning* (2000), pp. 359–366

4. R. Kohavi, G.H. John, Wrappers for feature subset selection. Artif. Intell. **97**(1–2), 273–324 (1997)

5. M. Modrzejewski, Feature selection using rough sets theory, in *Proceedings of the European Conference on Machine Learning* (1993), pp. 213–226

6. K.C.A. Wong, Simultaneous pattern and data clustering for pattern cluster analysis. IEEE Trans. Knowl. Data Eng. **20**(7), 911–923 (2008)

7. B.D. Fulcher, N.S. Jones, Highly comparative feature-based time-series classification. IEEE Trans. Knowl. Data Eng. **26**(12), 3026–3037 (2014)

8. C. Silva, B. Ribeiro, Margin-based active learning and background knowledge in text mining, in *Proceedings of the 4th International Conference of Hybrid Intelligent System* (2004). 0-7695-2291-2/04

9. C.C. Chen, M.C. Chen, TSCAN: a content anatomy approach to temporal topic summarization. IEEE Trans. Knowl. Data Eng. **24**(1), 170–183 (2012)

10. C. Clifton, R. Cooley, TopCat: data mining for topic identification in a text corpus (2000)

11. H. Frigui, O. Nasraoui, Simultaneous clustering and dynamic keyword weighting for Text documents (2004)

12. I. Al-Jadir, K.W. Wong, C.C. Fung, H. Xie, Text document clustering using memetic feature selection, in *Proceedings of the 9th International Conference on Machine Learning and Computing* (2017), pp. 415–420

13. I. Dhillon, J. Kogan, C. Nicholas, Feature selection and document clustering (2004)

14. J. Liang, F. Wang, C. Dang, Y. Qian, A group incremental approach to feature selection applying rough set technique. IEEE Trans. Knowl. Data Eng. **26**(2), 294–308 (2014)

15. J.C. Reynar, A. Ratnaparkhi, A maximum entropy approach to identifying sentence boundaries, in *Proceedings of the Conference on Applied Natural Language* (1994)

16. L. Nie, Y.-L. Zhao, M. Akbari, J. Shen, T.-S. Chua, Bridging the vocabulary gap between health seekers and healthcare knowledge. IEEE Trans. Knowl. Data Eng. **27**(2), 396–409 (2015)

17. M. Castellanos, HotMiner: discovering hot topics from dirty text. Surv. Text Min., 123–157 (2004)

18. N.K. Nagwani, A comment on the similarity measure for text classification and clustering. IEEE Trans. Knowl. Data Eng. **29**(9), 2589–2590 (2015)

19. N. Zhong, Y. Li, S.-T. Wu, Effective pattern discovery for text mining. IEEE Trans. Knowl. Data Eng. **24**(1), 30–44 (2012)

20. A. Christy, G. Meeragandhi, Combining bitemporal conceptual datamodel with multiway join relations for forecasting. Procedia Comput. Sci. **57**, 1104–1114 (2015)

21. A. Christy, M.G. Gandhi, S. Vaithyasubramanian, Cluster based outlier detection algorithm for healthcare data. Procedia Comput. Sci. **50**, 209–215 (2015)

Author information

Authors and Affiliations

School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

A. Christy & G. Meera Gandhi

Corresponding author

Correspondence to A. Christy .

Editor information

Editors and Affiliations

Department of Computer Application, SMIT, Sikkim Manipal University, Sikkim, India

Dr. Samarjeet Borah

Department of Automatics and Applied Software at the Faculty of Engineering, Aurel

Vlaicu University of Arad, Arad, Romania

Prof. Valentina Emilia Balas

Faculty of Technical Sciences, Jan Wyzykowski University, Polkowice, Poland

Dr. Zdzislaw Polkowski

Rights and permissions

Reprints and Permissions

Copyright information

© 2020 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Christy, A., Gandhi, G.M. (2020). Feature Selection and Clustering of Documents Using Random Feature Set Generation Technique. In: Borah, S., Emilia Balas, V., Polkowski, Z. (eds) Advances in Data Science and Management. Lecture Notes on Data Engineering and Communications Technologies, vol 37. Springer, Singapore. https://doi.org/10.1007/978-981-15-0978-0_6

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-0978-0_6

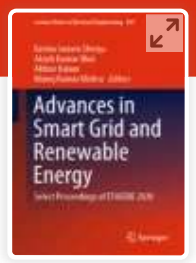
Published	Publisher Name	Print ISBN
14 January 2020	Springer, Singapore	978-981-15-0977-3

Online ISBN	eBook Packages
978-981-15-0978-0	Intelligent Technologies and Robotics

Intelligent Technologies and Robotics (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 625–634

Face Recognition Based Attendance System and Emotion Classification

Vallu Sri Satya Kala, Vanka Bhavyasri & S. Vigneshwari

Conference paper | First Online: 05 January 2021

306 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Attendance monitoring system is basic in all associations for checking the attendance of student or employee. In every association, attendance is taken physically by calling their register numbers or names and noted in their registers given by the association heads, and it consumes more time and effort. Face recognition is an example of recognition strategy and one of the most significant biometrics; it is utilized in an expansive range of utilizations. In our project, monitoring the attendance and classification of emotion is

simplified by face recognition mechanism by using MATLAB and sending SMS to absentees by using GSM mechanism. Attendance is recorded in the proposed work with the aid of using a digital camera that is connected in the study room or working environment. This constantly shoots the pictures of student or employee and detects his/her face. It classifies the faces in pixels by PCA algorithm and examines the detected faces with the database and marks the attendance. Without physical connection of individual, it gives the participation of present student or employee in the association. For face recognition, the eigenface algorithm is utilized and classification of emotions is finished by Viola–Jones algorithm. In our work, when a face of an understudy matches with the dataset, it denotes the understudy as present and afterward at last it will send the SMS to the versatile which understudies is missing, via GPRS/GSM modem and Arduino. This system also detects the faces and classifies emotion whether interested or uninterested and loads the data in to the excel sheet.

Keywords

Attendance

Principal component analysis (PCA)

Eigenfaces **Viola–Jones** **Arduino** **GSM**

This is a preview of subscription content, access via your institution.

Chapter **EUR 29.95**

Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_63
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook **EUR 234.33**

Price includes VAT (India)

- ISBN: 978-981-15-7511-2
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book **EUR 279.99**

Price excludes VAT (India)

- ISBN: 978-981-15-7714-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book **EUR 279.99**

Price excludes VAT (India)

- ISBN: 978-981-15-7510-5
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. S.K. Bhattacharyya, K. Rahul, Face recognition by linear discriminant analysis. Int. J. Commun. Netw. Secur. **2**(2), 2231–1882 (2013)

2. S. Marlin, S.S. Jebaseelan, B. Padmanabhan, G. Nagarajan, Power quality improvement for thirty bus system using UPFC and TCSC. Indian J. Sci. Technol. **7**(9), 1316–1320 (2014)

3. U.A. Kamerikar, M.S. Chavan, Experimental assessment of LDA and KLDA for face recognition. Int. J. Adv. Res. Comput. Sci. Manag. Stud. **2**(2), 137–146

4. A. Pravin, T. Prem Jacob, P. Asha, Enhancement of plant monitoring using Iot. Int. J. Eng. Technol. (UAE) **7**(3), 53–55 (2018)

5. M.P. Gawande, D.G. Agrawal, Face recognition using PCA and different distance classifiers. IOSR J. Electron. Commun. Eng. **9**(1), 1–5

6. K.S. Kumar, T. Sasikala, A Technique for web security using mutual authentication and clicking-cropping based image Captcha

technology. Int. Rev. Comput. Softw. **9**(1), 110–118 (2014)

7. N.N. Priyanka, N. Geetha, A.V.A. Mary, Cross-platform recognition of unknown identical users in multiple social media networks (2006)

8. N.H. Barnouti, Improve face recognition rate using different image pre-processing techniques, Am J Eng Re (AJER) **5**(4), 46–53 (2016)

9. K. Pradeep, T.P. Jacob, A hybrid approach for task scheduling using the cuckoo and harmony search in cloud computing environment. Wirel. Pers. Commun. **101**(4), 2287–2311 (2018)

10. R. Dheepa, D.U. Nandini, Privacy protection using sensitive data protection algorithm in frequent itemset mining of medical datasets. Res. J. Pharm. Biol. Chem. Sci. **7**(4), 308–316 (2016)

11. G. Nithya, S.L. Shabu, A novel framework in reusing the ontological health record. Res. J. Pharm. Biol. Chem. Scie. **7**(3), 215–220 (2016)

12. G. Nagarajan, R.I. Minu, A.J. Devi, Optimal nonparametric bayesian model-based multimodal BoVW creation using multilayer

pLSA. Circuits Syst. Signal Process. **39**(2), 1123–1132 (2020)

13. P.S. Maran, P.M. Velumurugan, B.P.D. Batvari, Wind characteristics and weibull parameter analysis to predict wind power potential along the South-east coastline of Tamil Nadu, in *International Conference On Intelligent Information Technologies* (Springer, Singapore, 2018), pp. 190–199

14. M. Divya, M.D. Kamalesh, Recovery of watermarked image from geometrics attacks using effective histogram shape based index. Indian J. Sci. Technol. **9**(44), 1–6 (2016)

15. Mohamed, B.K., Raghu, C., Fingerprint attendance system for classroom needs, in *India Conference (Indicon), 2012 Annual IEEE*. IEEE (2012), pp. 433–438

16. T. Lim, S. Sim, M. Mansor, RFID based attendance system, in *IEEE Symposium on Industrial Electronics & Applications*, 2009. ISIEA 2009, vol. 2. IEEE (2009), pp. 778–782

17. S. Kadry, K. Smaili, A design and implementation of a wireless iris recognition attendance management system, Information

Technology and Control, vol. 36, No. 3 (2007), pp. 323–329

18. V. Shehu, A. Dika, Using real time computer vision algorithms in automatic attendance management systems, Information Technology Interfaces (ITI), 2010 32nd International Conference (2010), pp. 397–402

19. R. Nithya, K. Muthu Priya, S. Vigneshwari, Stitching large images by enhancing surf and Ransac algorithm, in *2017 Second IEEE International Conference On Electrical, Computer And Communication Technologies*, 22–24 Febr 2017, SVS Engineering College, Coimbatore. Scopus. https://doi.org/10.1109/Icecct.2017.8117868

20. R. Vijai Chandra Prasad, M. Yashwanth Sai, P.R. Niveditha, T. Sasipraba, S. Vigneswari, S. Gowri, Low cost automated facial recognition system, in *2017 Second IEEE International Conference On Electrical, Computer and Communication Technologies*, 22–24 Febr 2017, SVS Engineering College, Coimbatore. 10.1109/Icecct.2017.8117829, Scopus

21. A. Sankari, S. Vigneshwari, Automatic tumour segmentation using CNN, in *Third IEEE International Conference On Science Technology, Engineering Management-*

ICONSTEM 2017, 23–24 Mar 2017, Chennai. Scopus (2017)

22. A. Singh, S.K. Singh, S. Tiwari, Comparison of face recognition algorithms on dummy faces. Int. J. Multimed. Its Appl. **4**(4), 121–135 (2012)

23. N.H. Barnouti, Face recognition using Eigen-face implemented on DSP processor. Int. J. Eng. Res. General Sci. **4**(2), 107–113

24. A. Mary Posonia, S. Vigneshwari, J. Albert Mayan, D. Jamunarani, Service direct: platform that incorporates service providers and consumers directly. Int. J. Eng. Adv. Technol. **8**(6), 3301–3304 (2019)

25. S. Prince Mary, B. Bharathi, S. Vigneshwari, R. Sathyabama, Neural computation based general disease prediction model. Int. J. Recent Technol. Eng. (IJRTE) 8(2), 5646–5449 (2019). ISSN: 2277–3878

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Vallu Sri Satya Kala, Vanka Bhavyasri & S. Vigneshwari

Corresponding author

Correspondence to Vallu Sri Satya Kala .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Kala, V.S.S., Bhavyasri, V., Vigneshwari, S. (2021). Face Recognition Based Attendance System and Emotion Classification. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra,

M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore.
https://doi.org/10.1007/978-981-15-7511-2_63

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7511-2_63

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

ScienceDirect

Materials Today: Proceedings

Volume 23, Part 1, 2020, Pages 34-38

Fabrication of p-type cubic γ-CuI by solid iodination process for energy conversion and storage applications

Karthik kumar Chinnakutti [a], Vengatesh Panneerselvam [a], Durai Govindarajan [a], Shyju Thankaraj Salammal [a, b]

Show more

Outline | Share | Cite

https://doi.org/10.1016/j.matpr.2019.05.387

Get rights and content

Abstract

A simple solid iodination method is used to achieve an efficient p-type cubic γ-CuI thin film from Cu_3N thin film. The crystallinity and crystal structure of γ-CuI thin films were confirmed by X-ray diffraction (XRD) technique. The maximum specific capacitance of CuI films was found to be 54 mF cm^{-2} at scan rate of 10 mV s^{-1} with excellent cycling stability. The optoelectronic behaviour was studied using UV–Vis–NIR spectroscopy and Hall measurements. These findings suggest that CuI film as a new inexpensive p-type material for third generation solar cells and good electrode for electrochemical energy storage devices.

Previous Next

Keywords

Solid iodination; p-type; γ-CuI; Specific capacitance; Hall measurements

Special issue articles Recommended articles

Cited by (3)

The transparent pn junction of CuI/SnO<inf>2</inf>/ZnO core-shell orderly nanoarrays towards photovoltaic conversion enhancement

2022, Materials Today Communications

Show abstract

The Transparent Pn Junction of Cui/Sno2/Zno Core-Shell Orderly Nanoarrays Towards Photovoltaic Conversion Enhancement

2022, SSRN

The Transparent Pn Junction of Cui/Sno2/Cu2o Core-Shell Orderly Nanoarrays Towards Photovoltaic Conversion Enhancement

2022, SSRN

© 2019 Elsevier Ltd. Peer-review under responsibility of the scientific committee of the Advanced Materials for Clean Energy and Health Applications (AMCEHA).



Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

RELX™

International Conference on Communication and Signal Processing, July 28 - 30, 2020, India

Eyeball based Cursor Movement Control

Sivasangari.A, Deepa.D, Anandhi.T, Anitha Ponraj and Roobini.M.S

***Abstract*—An individual Human computer interference system is being introduced. In olden times, as an input device the mouse and keyboard were used by human computer interference system. Those people who are suffering from certain disease or illness cannot be able to operate computers. The idea of controlling the computers with the eyes will serve a great use for handicapped and disabled person. Also this type of control will eliminate the help required by other person to handle the computer. This measure will be the most useful for the person who is without hands through which they can operate with the help of their eye movements. The movement of the cursor is directly associated with the center of the pupil. Hence our first step would be detecting the center of point pupil. This process of pupil detection is implemented using the Raspberry Pi and OpenCV. The Raspberry Pi has a SD/MMC card slot which is used for placing the SD card. The SD card is boosted with the operating system that is required starting up of Raspberry Pi. The Raspberry PI will get executed once the application program is loaded into it.**

***Index Terms*—Human Computer Interaction (HCI), Eyeball movement, Computer, OpenCV, Support Vector Machine.**

I. INTRODUCTION

AS the computer technologies are growing rapidly, the importance of human computer interaction becomes highly notable. Some persons who are disabled cannot be able to use the computers. Eye ball movement control mainly used for disabled people. Incorporating this eye controlling system with the computers will make them to work without the help of other individual. Human-Computer Interface (HCI) is focused on use of computer technology to provide interface between the computer and the human. There is a need for finding the suitable technology that makes the effective communication between human and computer. Human computer interaction plays the important role .Thus there is a need to find a method that spreads an alternate way for making communication between the human and computer to the individuals those who have impairments and give them an equivalent space to be an element of Information Society [1-5].

In recent years, the human computer interfaces are attracting the attention of various researchers across the globe. Human computer interface is an implementation of the vision-based system for eye movement detection for the disabled people.

In the proposed system, we have included the face detection, face tracking, eye detection and interpretation of a sequence of eye blinks in real time for controlling a non-intrusive human computer interface. Conventional method of interaction with the computer with the mouse is replaced with the human eye movements. This technique will help the paralyzed person, physically challenged people especially person without hands to compute efficiently and with the ease of use. Firstly, camera captures the image and focuses on the eye in the image using OpenCV code for pupil detection. This results the center position of the human eye (pupil). Then the center position of the pupil is taken as a reference and based on that the human or the user will control the cursor by moving left and right [6-9].

This paper organization is described as follows. Section II describes existing solutions to find the cursor movement using some 3D models. In Section III we present how cursor is working based only on Eyeball movement using OpenCV methodology. In Section IV how the cursor is moving using eyeball with example with the better solutions. And the Conclusion part are presented in section V.

II. RELATED WORK

The basic actions of a mouse are mouse click and mouse movement. The advance technology replaces this mouse movement by eye motion with the help of an OpenCV. The mouse button click is implemented by any of the facial expressions such as blinking eyes, opening mouth and head movement. This model introduces a novel camera mouse driven by 3D model based bias face tracking technique. In personal computer(PC) due to the standard configuration it achieves human machine interaction through faster visual face tracking and provides a feasible solution to hand-free control. The face tracker used here is based on 3D model to control the mouse and carry out mouse operations.

Gaze estimation can be used in Head-mounted display (HMD) environments since they can afford important natural computer interface cues. This new gaze estimation is based on 3D analysis of human eye. There are various commercial products which use gaze detection technology. In this method, the user has to point only one point for calibration it will then estimate the gaze points. The facial features such as eyes and nose tip are recognized and tracked to avoid the traditional mouse movements with the human face for human interaction with the computer. This method can be applied to face scales in a wide range.

Sivasangari.A, Deepa.D, Anandhi.T, Anitha.Ponraj and Roobini.M.S are with the Sathyabama Institute of science and technology, Chennai. (email: sivasangarikavya@gmail.com deepa21me@gmail.com anandhi.cse@sathyabama.ac.in anisainosoft@gmail.com roobinims@gmail.com)

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:06:37 UTC from IEEE Xplore. Restrictions apply.

Six-Segmented Rectangular (SSR) filter and support vector machine are used for fast extraction of face candidates and face verification respectively. This comprises of our basic strategy for detection. Using JAVA(J2ME) for face candidate detection, the scale adaptive face detection and tracking system are implemented to perform left/right mouse click events when the left/right eye blinks. Camera mouse has been used for disabled people to make an interaction with the computer. [1]

The camera mouse is used to change all roles of traditional mouse and keyboard actions. The proposed system can give all mouse click events and keyboard functions. In this method, the camera mouse system along with the timer acts as left click event and blinking as right click event. The real time eye-gaze estimation system is used for eye controlled mouse for assisting the disabled. This system based on the methodology in which a general low-resolution webcam is used, but detects the eyes and track gaze accurately in less expense and without specific equipment. PIR sensor is specifically used for the human movement detection. This paper introduces a novel camera mouse driven by visual face tracking based on a 3D mode. This camera has a standard configuration for PCs with increased computation speed and also providing feasible solution to hands free Control through visual face tracking. Human facial expressions can be classified as rigid motion and non rigid motions. The rigid motions are rotation and translation whereas the non-rigid motions are opening, closing and stretching of the mouth.

Firstly, we use a virtual eyeball model which is based on the 3D characteristics of the human eyeball. By using a camera and three collimated IR-LEDs secondly, we will be calculating the 3D position of virtual eyeball and gaze vector [2]. Thirdly the calculation of 3D eye position and gaze position on a HMD monitor is allowed. This used simplified complex 3D converting calculations that have three reference frames (the camera, the monitor and the eye reference frames). Fourth, based on kappa compensation, a simple user-dependent calibration method was proposed by gazing at one position.

In our work, we are trying to compensate the need of people who have hand disabilities and could not be able to use computer resources without other individual's help. Our application mainly uses facial features to interact with the computer hence there will be no need of hands to operate the mouse [3].

Paralysis is a special case, in which the loss of muscle functions in part of your body. It happens when something goes wrong with the way messages pass between your brain and muscles. When such a thing happens, the person's ability to control movement is limited to the muscles around the eyes [4]. Blinking and movement of eyes is the only way of communication for them. To such communication defects the assistance given is often intrusive that is it requires a special hardware or a device. The alternate way for interfacing is through a non-intrusive communication system such as Eye Keys which works without special lightning. The eye direction is detected if the person looks at the camera, which can be used to control various applications [5].

III. Proposed Work

The proposed system uses Raspberry Pi board of version 3, which is attached with the Monitor, PIR Sensor, and Camera. These materials are attached by USB adaptors. Raspberry plays a vital role in the working module that keeps the eye movement with sensors. Raspberry pi uses SD card, to install raspbian is along with programming codes. PIR sensor also used for detecting human movement.

A. Face Detection

The computer technology which is used for a variety of applications by identifying the human faces in digital images is called as face detection. The proposed method detects features from the face. A simple face tracking system was developed. Face images can be analyzed without ever requiring any interaction with the user/person. Facial recognition can be used as an important measure of tracking the attendance and time information. Human face provide facial information that can be used for many applications like emotion recognition and human computer interface. Local binary pattern algorithm can be used for feature extraction.3×3 pixel image can be taken from the web camera. Encoding operation can be performed pixel values and transformed in to binary value 0 or 1.The face image is divided into N blocks. The thresholding function is described below:

$$S(g_c, g_i) = \begin{cases} 1, if\ g_c \geq g_i \\ 0, if\ g_c < g_i \end{cases} \quad (1)$$

Weight values are calculated for every neighbour and calculate the LBP value.

B. Eye Region Detection

The exact position of the pupil is known by using vertical integral projection and horizontal projection. These projections divide the whole picture to homogenous subsets. The arbitrary threshold is used in the proposed method. The noise can be removed by using Gaussian filter. The strong pixel value is based on minimum gradient point. The lower threshold protects against splitting edges in the contrast region. Circular Hough transform is used for finding the inner and outer boundaries. Hough transform check all the edge points with center coordinates.

C. Eye Movement Classification

The different eye-motions are classified with the help of support vector machine classifier. The eye-movements are eye open, eye close, eyeball left and eyeball right are captured by web camera. SVM can analyze data and used for classification and regression analysis. SVM is a set of associated supervised learning functions used for classification and regression problems. In SVM, the multi class file is used. This PIR sensor is connected to the General Purpose I/O port of the raspbian board, it detects the eye pupil movement and hence makes the camera to start capturing the images. Sensor can cover the range up to 5cm.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:06:37 UTC from IEEE Xplore. Restrictions apply.

Once when the sensor detects the movement of eye pupil the camera starts to capture the images and send it into the raspbian board through an USB cable. The camera used here is an USB web camera of affordable cost. Then the transferred image is processed and monitored out. The function of the SD card is to store the Raspbian Jessie operating system module and it stores the program in it. The Raspberry Pi board is activated using Python programming language. The SD card capacity is up to 8GB.

The monitor input is got from the HDMI port of the raspberry board. HDMI is a High Definition Multimedia Interface port which is used to monitor the uncompressed video data. The HDMI converts digital image signal into analog signal and gives it to the monitor. Camera is used to capture the movement of eye. The center portion of the eye as well as the position of the people is identified for getting the various movements of eye. The process is implemented using Raspberry Pi. In Fig. 1 shows that the pupil reference has the coordinates of (x, y). Raspberry pi will be combining with USB Camera. Raspberry pi will be use SD card, then the install raspbian OS and Opencv on raspberry pi. First image will be capture by USB Camera. Focus on eye in image and detect the center position of pupil by Opencv code.

Support vector machine algorithm is used for classification. Eye image is denoted as a vector I which represents the pixel values. Training images are divided into two sets. The task of SVM is to determine the new data belongs to either positive set or negative set. Principle component analyses are applied to the training data and reduce the dimensionality of data.

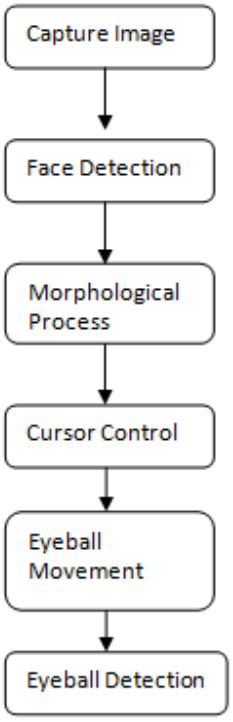


Fig. 1. Flow of Proposed Work

IV. Performance Analysis

After initiating the OS the operation starts. At first PIR sensor is used to detect the presence of individual in front of system. If the person is detected then camera will ON, where image will be captured by USB Camera. Focus on eye in image is shown in Fig. 2. Now the center position of pupil is detected in OpenCV. Take the exact position of pupil as reference, and then the next the different value of X, Y coordinates will be set for accurate command.

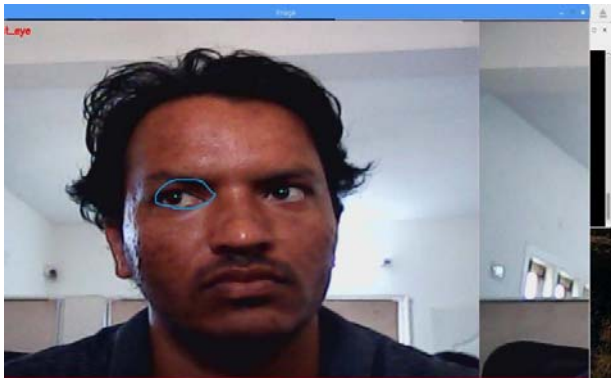


Fig. 2. Detection of Eye

$$d = \frac{max\left\{\left||L-L'|\right|_2, \left||R-R\right|\right.}{L-R} \quad (2)$$

Detection error is the one of the parameter to evaluate the accuracy of proposed work. Images can be collected from the BioID database. The proposed method achieves better accuracy than existing methods. The following Fig. 3 shows the efficiency at different error levels.

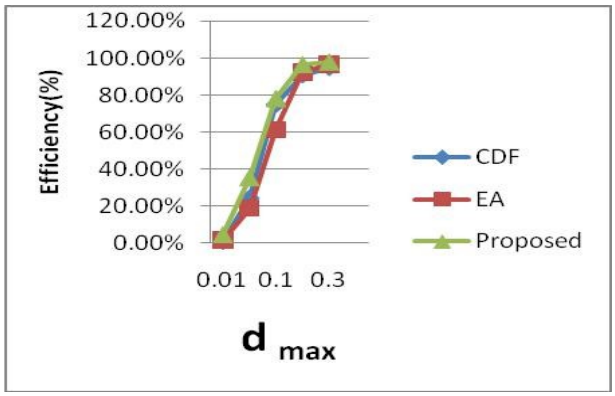


Fig. 3. Efficiency Analysis

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:06:37 UTC from IEEE Xplore. Restrictions apply.

V. Conclusion

From the process implemented it is cleared that the cursor can be controlled by the eyeball movement i.e., without using hands on the computers. This will be helpful for the people having disability in using the physical parts of the computers to control the cursor points. Because the cursor points can be operated by moving the eyeballs. Without the help of others disabled people can use the computers. This technology can be enhanced in the future by inventing more techniques like clicking events as well as to do all the mouse movements and also for human interface systems using eye blinks. Technology also extended to the eyeball movement and eye blinking to get the efficient and accurate movement.

References

[1] Jilin Tu, Thomas Huang, Elect and Comp EngrDept, Hai Tao, ElectEnggDept, "Face as Mouse through Visual Face Tracking",IEEE,(2005).

[2] EniChul Lee Kang Ryoung Park "A robust eye gaze tracking methodbased on a virtual eyeball model", Springer, pp.319-337, Apr (2008).

[3] John J. Magee, MargritBetke, James Gips, Matthew R. Scott, and Benjamin N.Waber"A Human-Computer Interface Using Symmetry Between Eyes to Detect Gaze Direction" IEEE Trans, Vol. 38, no.6,pp.1248-1259, Nov (2008).

[4] SunitaBarve, DhavalDholakiya, Shashank Gupta, DhananjayDhatrak, "Facial Feature Based Method For Real Time Face Detection and Tracking I-CURSOR", International Journal of EnggResearchand App., Vol. 2, pp. 1406-1410, Apr (2012).

[5] Yu-Tzu Lin Ruei-Yan Lin Yu-Chih Lin Greg C Lee "Real-time eye-gaze estimation using a low-resolution webcam", Springer, pp.543-568, Aug (2012).

[6] Samuel Epstein-Eric MissimerMargritBetke "Using Kernels for avideo-based mouse-replacement interface", Springer link, Nov (2012)

[7] Hossain, Zakir, Md Maruf Hossain Shuvo, and Prionjit Sarker. "Hardware and software implementation of real time electrooculogram (EOG) acquisition system to control computer cursor with eyeball movement." In 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), pp. 132-137. IEEE, 2017.

[8] Lee, Jun-Seok, Kyung-hwa Yu, Sang-won Leigh, Jin-Yong Chung, and Sung-Goo Cho. "Method for controlling device on the basis of eyeball motion, and device therefor." U.S. Patent 9,864,429, issued January 9, 2018.

[9] Lee, Po-Lei, Jyun-Jie Sie, Yu-Ju Liu, Chi-Hsun Wu, Ming-Huan Lee, Chih-Hung Shu, Po-Hung Li, Chia-Wei Sun, and Kuo-Kai Shyu. "An SSVEP-actuated brain computer interface using phase-tagged flickering sequences: a cursor system." Annals of biomedical engineering 38, no. 7 (2010): 2383-2397

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:06:37 UTC from IEEE Xplore. Restrictions apply.

ScienceDirect

Biocatalysis and Agricultural Biotechnology
Volume 27, August 2020, 101655

Extracellular synthesis of nanoselenium from fresh water bacteria *Bacillus sp.*, and its validation of antibacterial and cytotoxic potential

S. Bharathi [a], S. Kumaran [b, c], G. Suresh [a], M. Ramesh [d], V. Thangamani [a], S.R. Pugazhvendan [e, f], K. Sathiyamurthy [g]

Show more

Share Cite

https://doi.org/10.1016/j.bcab.2020.101655 Get rights and content

Abstract

In this present study, selenium reducing bacteria were isolated from fresh water pond. The cell free supernatant from B2 strain was found to effectively reduce sodium selenite and was identified as Bacillus *sp.* based on 16S r RNA sequencing. The selenium nanoparticles (Se NPs) synthesised using *Bacillus sp.* B2 were characterised by UV- spec, XRD, FTIR, FESEM and EDX. The selenium nanoparticles exhibited maximum peak at 280 nm in UV spectrophotometric analysis due to the surface plasmon resonance. FTIR analysis of cell free supernatant from *Bacillus sp.* B2 and synthesised selenium nanoparticles reveals that the proteins were mainly responsible for the reduction and stabilisation of Se nanoparticles. The XRD pattern shows the synthesised Se NPs were crystalline in nature with the average size of 33.2 nm. The SEM analysis reveals the Se NPs are spherical in shape with the size range of 20-50 nm and EDX pattern shows selenium exist as major element. The antibacterial activity of the Se NPs shows activity against all the tested bacteria. The lowest MIC and MBC of Se NPs was 8 μg/200 μL against Pseudomonas aeruginosa (NCIM 5031). The cytotoxicity of SeNPs against Human non-small lung cancer cell line A549 was tested by MTT assay that showed IC50 value of 3 μg/mL. The ET/AO staining of the Se NPs treated cells further revealed the result. Based on this we conclude that fresh water bacteria *Bacillus sp.* B2 can be used for selenium nanoparticles synthesis which is an ecofriendly and easy process for biomedical applications.

Introduction

Selenium is one of the essential trace elements in the body due to anti-oxidant as well as pro oxidative effects and has great importance in medicine (Ingole et al., 2010). It is important for the synthesis of nearly 13 proteins and acts as cofactor for various enzymes like glutathione peroxidase and thioredoxin reductases (Afzal et al., 2019). In the recent year bright red colored selenium nanoparticles with unique biological activity and excellent properties have drawn worldwide attention (Bai et al., 2017). The high biological activity and low toxicity of selenium nanoparticles made them unique to be used in medicine (Fesharaki et al., 2010).

Se NPs found to possess excellent anticancer and free radical scavenging activity. Se NPs used in many disease conditions including cancer, diabetes, inflammatory disorders, liver fibrosis, infectious disease and drug induced toxicities (Khurana et al., 2019). Se NPs are mostly synthesised by the methods employing elevated temperature and high pressure and are highly dangerous to environment (Torres et al., 2012). The application of selenium nanoparticles is limited due to the lack of economic and ecofriendly methods to synthesise selenium nanoparticles (Bai et al., 2017). The biologically synthesised nanoparticles are cost-effective, ecofriendly, less toxic, and also results in stable nanoparticles without aggregation (https://www.ncbi.nlm.nih.gov/pubmed/?term=Shah%20M%5BAuthor%5D&cauthor=true&cauthor_uid=28793638 Shah et al., 2015).

Many bacteria, fungi and plants were reported for the selenium nanoparticles synthesis. Different microorganisms can reduce metal ions into elemental nanoparticles. The metal ions get absorbed into cells and reduction mostly happens in the intracellular space of

microorganisms. The release of nanoparticles from the cell needs appropriate extraction methods which use liquid nitrogen and solvents for extraction that are laborious and time consuming (Zare et al., 2013). There has been great interest in use of cell free supernatants for the reduction of metal ions (Shahverdi et al., 2007). There will be no expensive treatments for cell disruption method to separate the synthesised nanoparticles. The nanoparticles' activity is highly dependent on shape, size and capping agents hence, exploitation of bacteria from less explored ecosystems is the need of the hour. Hence the present study aims for the extra cellular synthesis of selenium nanoparticles from bacteria of fresh water ecosystems.

Section snippets

Isolation of bacteria from fresh water sample

Water samples were collected from five different locations within the Thadhasamudhram pond present in Ayyangarkulam village (12.7843° N, 79.6628° E), near Kanchipuram, Tamil Nadu using sterile plastic container. Isolation of bacteria was performed by serial dilution followed by plating in nutrient agar plates and nutrient agar supplemented with 1 mM sodium selentie (Mishra et al., 2011). After the incubation for 24 h at 28 °C, total numbers of heterotrophic and selenite-reducing bacteria were...

Isolation of selenium reducing bacteria from fresh water sample

Microorganisms from aquatic environment have been less exploited as compared to terrestrial ecosystem. Hence in this study, fresh water bacteria were exploited for the biotransformation of selenate and selenite oxoanions to elemental selenium with the formation of selenium nanoparticles. The collected water sample were analysed for heterotrophic bacteria and found to be 7.5×10^6 CFU/mL and selenite reducing bacteria with 6×10^3 CFU/mL. The analysis of 16S rRNA gene sequence of the selected...

Conclusion

In the current investigation, the bacteria isolated from fresh water ecosystem were screened for extracellular synthesis of selenium nanoparticles. The potent Se NPs producing bacterial strain B2 was identified as *Bacillus sp.* by 16s r RNA sequencing. The characterization studies reveal that various organic molecules in the cell free supernatant from bacterial strain B2 responsible for selenite reduction. The Se NPs was amorphous, spherical shaped with the average diameter of 20–50 nm. Thus the ...

CRediT authorship contribution statement

S. Bharathi: Methodology, Writing - original draft, Writing - review & editing. **S. Kumaran:** Formal analysis. **G. Suresh:** Formal analysis. **M. Ramesh:** Data curation. **V. Thangamani:** Data curation. **S.R. Pugazhvendan:** Formal analysis. **K. Sathiyamurthy:** Formal analysis....

Declaration of competing interest

Authors have no conflict of interest to declare....

Acknowledgement

Authors are thankful to Dr. K. R. Venkatesan, Principal, Sri Sankara Arts and Science College for providing lab facilities....

Research data for this article

Data not available / No data was used for the research described in the article

Further information on research data

References (37)

C. Xu *et al.*
Preparation, characteristics and antioxidant activity of polysaccharides and proteins-capped selenium nanoparticles synthesized by Lactobacillus casei ATCC 393
Carbohydr. Polym. (2018)

A.R. Shahverdi *et al.*
Rapid synthesis of silver nanoparticles using culture supernatants of Enterobacteria: a novel biological approach
Process Biochem. (2007)

R.R. Mishra *et al.*
Reduction of selenite to red elemental selenium by moderately halotolerant *Bacillus megaterium* strains isolated from Bhitarkanika mangrove soil and characterization of reduced product
Chemosphere (2011)

A.J. Kora *et al.*
Biomimetic synthesis of selenium nanoparticles by *Pseudomonas aeruginosa* ATCC 27853: an approach for conversion of selenite
J. Environ. Manag. (2016)

A. Khurana *et al.*
Therapeutic applications of selenium nanoparticles
Biomed. Pharmacother. (2019)

M. Jeyaraj *et al.*
Biogenic silver nanoparticles for cancer treatment: an experimental report
Colloids Surf. B Biointerfaces (2013)

B. Afzal *et al.*
Screening of cyanobacterial strains for the selenium nanoparticles synthesis and their anti-oxidant activity
Biocatalysis and Agricultural Biotechnology (2019)

V. Alagesan *et al.*
Green synthesis of selenium nanoparticle using leaves extract of *Withania somnifera* and its biological applications and photocatalytic activities
Bio Nano Science (2019)

A. Ananth *et al.*
Synthesis and characterization of nano-selenium and its antibacterial response on some important human pathogens
Curr. Sci. (2019)

K. Bai *et al.*
Preparation and antioxidant properties of selenium nanoparticles-loaded chitosan microsphere
Int. J. Nanomed. (2017)

View more references

Cited by (15)

Selenite bioreduction and biosynthesis of selenium nanoparticles by Bacillus paramycoides SP3 isolated from coal mine overburden leachate

2021, Environmental Pollution

Show abstract

Biosynthesis of Se-Nanorods using Gum Arabic (GA) and investigation of their photocatalytic and cytotoxicity effects

2021, Inorganic Chemistry Communications

Show abstract

Interaction of nanomaterials with microbes

2021, Biodegradation and Biodeterioration at the Nanoscale

Show abstract

Synthesis of selenium nanoparticles by using microorganisms and agri-based products

2021, Agri-Waste and Microbes for Production of Sustainable Nanomaterials

Show abstract

Synthesis of nanoparticles using microorganisms and their applications: a review

2022, Environmental Chemistry Letters

Selenium nanoparticles: A review on synthesis and biomedical applications

2022, Materials Advances

View all citing articles on Scopus

Recommended articles (6)

Research article

Efficacy of Biogenic Selenium Nanoparticles from an Extract of Ginger towards Evaluation on Anti-Microbial and Anti-Oxidant Activities

Colloid and Interface Science Communications, Volume 29, 2019, pp. 1-8

Show abstract

Research article

Immobilization of elemental mercury by biogenic Se nanoparticles in soils of varying salinity

Science of The Total Environment, Volume 668, 2019, pp. 303-309

Show abstract

Research article

Antimicrobial and antibiofilm effects of selenium nanoparticles on some foodborne pathogens

LWT - Food Science and Technology, Volume 63, Issue 2, 2015, pp. 1001-1007

Show abstract

Research article

Biosynthesis of selenium-nanoparticles and -nanorods as a product of selenite bioconversion by the aerobic bacterium *Rhodococcus aetherivorans* BCP1

New Biotechnology, Volume 41, 2018, pp. 1-8

Show abstract

Research article

Insights into selenite reduction and biogenesis of elemental selenium nanoparticles by two environmental isolates of *Burkholderia fungorum*

New Biotechnology, Volume 34, 2017, pp. 1-11

Show abstract

Research article

Biosynthesis of selenium nanoparticles mediated by fungus *Mariannaea* sp. HJ and their characterization

Colloids and Surfaces A: Physicochemical and Engineering Aspects, Volume 571, 2019, pp. 9-16

Show abstract

View full text

© 2020 Elsevier Ltd. All rights reserved.

Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

ELSEVIER

RELX™

Evaluation of Tweets for Content Analysis Using Machine Learning Models

Amit Pimpalkar
School of Computing,
Sathyabama Institute of Science and Technology,
Chennai, India
amit.pimpalkar@gmail.com

R. Jeberson Retna Raj
Associate Professor, Department of IT,
Sathyabama Institute of Science and Technology,
Chennai, India
jebersonretnaraj.it@sathyabama.ac.in

Abstract— **Now-a-days we are virtually residing in the digital networking era, and constantly sharing thoughts, opinions about things that are associated with us. As a result, social networking sites, especially Twitter, get flooded with tremendous information in the form of opinions. They are unstructured so disseminated, and a solid base has to be established such that they can be seen as useful knowledge about a specific issue. This work focuses on the evaluation of different models with respect to feature selected and dataset of user's feelings on twitter data using different pre-processing measures. The dataset is extracted from Kaggle and Twitter, pre-processing performed using NLTK, Scikit-learn and features selection, extraction done for a Bag of Words (BOW), Term Frequency (TF) and Inverse Document Frequency (IDF). For polarity identification and evaluation, five different Machine Learning (ML) algorithms were compared. The performance comparison of these algorithms in an attempt to decide which algorithm is ideally suited for the selected dataset in view of the recall, accuracy, F1-score, and precision observed. As the evaluation results, an SVM classifier outperformed the other implemented models.**

Keywords- Machine Learning; Sentiment Analysis; Bag of Words; TF-IDF; Evaluations Metrics; Twitter Classifier

I. INTRODUCTION

Twitter is a common social networking site where users connect through messages identified as 'tweets'. With around 6,000 tweets are tweeted every second and it has only a 140 character-limit. Tweets include misspelling, the dialect of slang and the imaginative definitions that lead to the serious challenges in data processing and analysis. Opinion Mining (OM) coincides with technologies which are capable of detecting positive or negative expressions. "What other people think??" is often affected people's decisions on each and every subject [5]. The immense amounts of information available on the Internet has caused numerous organizations to become involved in this, because making use of this knowledge from a future viewpoint would be a tremendous gain for them. It refers to a totally new and wide analysis area named Sentiment Analysis (SA). SA plays a major role here to aid in the quest for a hidden emotion or meaning within the message. Labeling a tweet as a positive or a negative word is normal for humans, but it is not sufficient to handle huge volumes of information on the internet. In order to solve these problems, ML strategies are introduced. If there are more positive words in the document, a positive polarity and vice versa is defined by ML algorithms.

In this article, in the sense of sentence-level emotion recognition, we compare the ML methods Multinomial Naive Bayes (MNB), Logistic Regression (LR), Decision Trees (DT), XGBoost (XGB) and Support Vector Machines (SVM). This research helps to examine the most suitable algorithm for the chosen algorithms for the analysis of SA Twitter results.

The remnant of the manuscript is organized as follows: In the "Literature Review" section, we provide related work done prior to this. The "System Design" section presents a system flow diagram for our research. In the "Experimental Evaluation" section, we carry out experiments to evaluate the performance of the classification algorithms. Finally, the "Conclusion and Future Scope" section conclude the paper with the guidelines for the future research on the continuation track.

II. LITERATURE REVIEW

Many researchers concentrated on the topic of classifying emotions. Several datasets are accessible publicly online and some scholars have started to examine the findings. As our research focused on the positive or negative evaluation, we concentrated our literature on classifying sentence level sentiments.

Through the Stanford Testing Sentiment data set [14], the efficacy of NB and SVM classifiers was evaluated. Application of NB and SVM was supported by authors of [2] in an evaluation of political opinions. In order to classify documents polarity, various ML approaches were contrasted, and the LR proposed the best possible accuracy [7]. Measured by NB, Max Entropy and SVM Classifier efficiencies [6], SVM was found to function better. Different approaches used in pre-processing typically in Twitter data sets have been compared [8]. The study [21] addressed the results in the effective extraction and selection of features of the pre-processing text structure and then classified tweets as positive, negative and neutral using ML techniques. A pre-processed data frame based on the NLP has been developed by [22] for the purpose of filtering messages, where authors have used a BoW model and a TF-IDF model with a SA depends on Twitter APIs. In the

978-1-7281-9393-9/20/$31.00 ©2020 IEEE
DOI: 10.1109/CICN.2020.81

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

authors' opinion, the introduction of diagrams and images greatly increases the description of emotions. They also demonstrated that the pre-processing and sample size method dramatically influence the relationship between consumer confidence and stock returns.

R. Othman [24] suggested that continuous word output as a classification function for SemEval usage of RNN-LSTM be correlated with a feel for writing. Three signal detection algorithms were supported by [10] for volume tweets, posts for sentiment, and top hashtags. The solution to calculating the importance of each research paper was suggested [27] using a process integrating TF-IDF and LDA scheme. A two-stage approach to ML algorithms and combinational fusion in data processing was formulated [18]. A word dependent N-gram method to break Twitter tweets randomly into plausible, not realistic tweets that [23] implements. Different studies have tested the effectiveness of deep CNN and RNN-based model neural networks utilizing embedded text vectors in SA with the specific tuning of hyper-parameters in a wide variety of researchers on film ratings, political commentary and on-line user reviews. [9-10 30-31, 33-36]

III. System Design

Fig 1 shows the data flow through numerous Twitter Content Processing Framework modules.

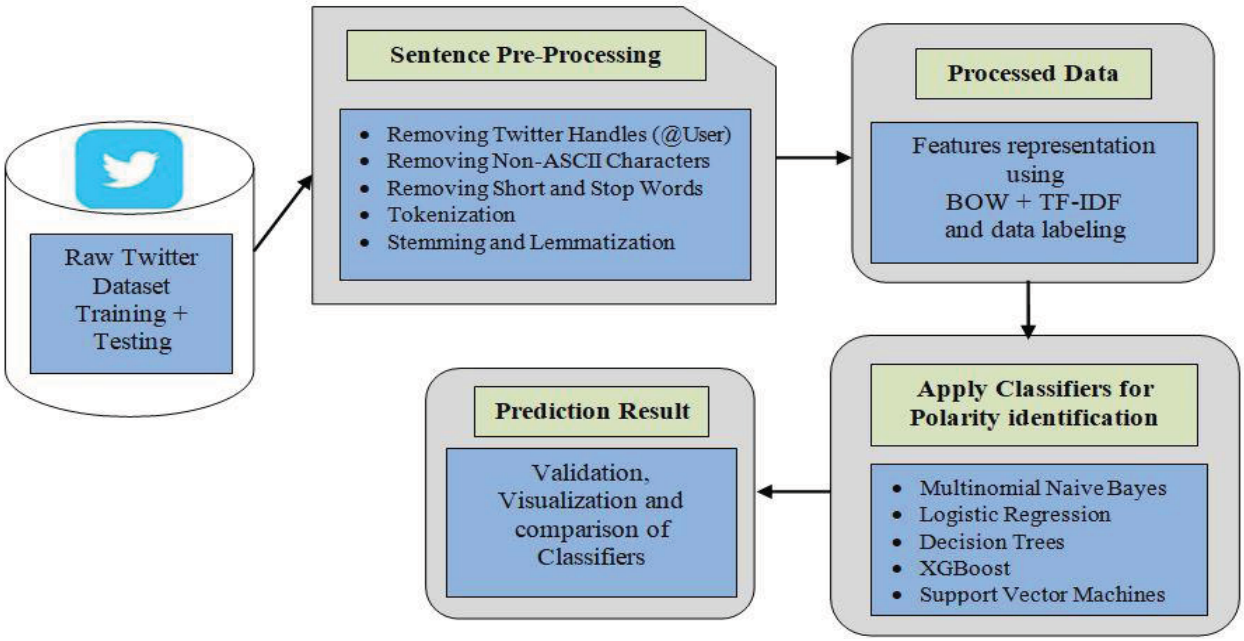


Fig 1: Twitter Content Processing Framework

A. Dataset

This research utilizes two data sets, Twitter reviews of common, domain-free tweets and the other of a movie domain: an Internet Film Database (IMDb). The former data set provides 49159 (positive and negative 45467) ratings, and the latter data collection comprises 50,000 (positive 25,000 and negatives 25,000). The data sets were downloaded from Kaggle, though twitter was crawled. By 67 and 33 percentages, our dataset is divided into 2 parts for training and testing. The goal is to acquire a classification model that can differentiate between positive and negative evaluations of various measurement parameters, most widely used for predicting the effectiveness of algorithms including precision, F1-score, recall, and precision.

B. Initial Text Mining and Pre-Processing

Pre-processing is one of the key measures to render reviews noiseless, for instance, unnecessary letters, symbols, gestures, etc, at the classification point, as the evaluation phrases might differentiate either two or three features. This could make it impossible for the paragraph to discern either positive or negative. With the help of the python regular expression function, we prepared the following.

- Removing Twitter Handles (@User)
- Removing Non-ASCII Characters
- Removing Short and Stop Words
- Tokenization
- Stemming and Lemmatization

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

C. Features Selection and Extraction in Sentiment Analysis:

Extrapolation includes dataset characteristics extraction in ML algorithm supported format. The following characteristics were incorporated and applied for the analysis in designing the concept for this test.

- **Terms presence and frequency:** It relates to the amount of occasions the terms occur in the texts, and is also called the feature n-gram.

$$\mathrm{TF(t)} = \frac{\text{(Number of times term t appears in a document)}}{\text{(Total number of terms in the document)}} \quad (1)$$

- **Document Frequency:** Frequency of Documents is the number of times features appear in all texts.
- **Term Frequency-Inverse Document Frequency (TF-IDF):** The TF-IDF weight used for the extraction of information in text mining. This weight is a quantitative metric that is used to determine how important a term in a list or corpus is to a document. The weight of the TF-IDF consists of two terms: The first is the normalized TF, where the majority of words occurs in a document is measured by the total number of terms in that text; the second is the IDF, which is determined as the natural log of documents in the collection distributed by the number of documents in which the expression exists. [1, 3-4, 11-13, 15-17, 24, 26, 29, 37].

$$\mathrm{IDF(t)} = \log\left(\frac{\text{Total number of documents}}{\text{Number of documents with term t in it}}\right) \quad (2)$$

$$\mathrm{TFIDF(t)} = \mathrm{TF(t)} * \mathrm{IDF\ (t)} \quad (3)$$

- **Bag of Words:** The feature selection techniques treat the documents either as a group of words or BOW. The idea is to enable a free textual content into a vector which can be used as an input or output for an ML model. [4, 15-16, 19-20, 24, 29, 32]. This is illustrated as follows with an example, considering, we have 3 documents as

D1: This movie is long but good.
D2: This movie is too slow and long
D3: This movie is awesome
Here we need to develop a vocabulary utilizing identical terms from all the texts by ignoring case and punctuation because they don't produce the necessary knowledge for the model as **[**'movie', 'long', 'good', 'slow', 'awesome']
Here, D=3, N=5
The matrix M of size 3 X 5 will be represented as:

TABLE 1: EXAMPLE OF DOCUMENT CONTAINING TERM FREQUENCIES OF WORD

	movie	long	good	slow	awesome
D1	1	1	1	0	0
D2	1	1	0	1	0
D3	1	0	0	0	1

The above Table 1 shows the training features in each document which contain term frequencies of each word as we used for our illustration.

D. Classification Algorithm

This segment reflects on the theoretical background of traditional ML algorithms used in this work.

- **Multinomial Naive Bayes**: MNB is a sophisticated Naive Bayes algorithm extension which is best suited for classification of text documents. While NB models a text as an inclusion of and absence of specific terms, MNB models the word counts directly and alters the categories to be used. It can be described as; MNB calculates the conditional probability of a given word given to a class as the relative frequency of term (t) in documents belonging to the class (c). The variation takes into consideration the amount of term (t) occurrences from class (c) in training documents, including several instances. The probability of a document d being in class c is computed as

$$P(c|d) = P(c)\prod_{1\le k\le n_d} P(t_k|c) \quad (4)$$

Where $P(t_k|c)$ is the conditional probability of term t_k occurring in a document of class c.
P(c) is the prior probability of a document occurring in class c.

- **Logistic Regression**: L is also one of the most useful computational methods as it is capable of transparently evaluated the value of each function. It collects, weights, removes sums and transfers the total to generate a likelihood via the sigmoid function in order to prevent appearing in LR. For decision-making, thresholds are used. Throughout LR, the linear approach is primarily used as a guide to another process in the following relationship, such as *g*

$$h_{\emptyset}(x) = g(\emptyset^T x) \quad (5)$$

where $0 \le h_{\emptyset} \le 1$. In this case *g* is the logistic or sigmoid function that can be given as follows

$$g(z) = \frac{1}{1+e^{-z}} \quad (6)$$

- **Decision Trees**: This classification method is supervised and engages data to construct decision trees with identifying labels and the framework for the test results. When each tree node represents the data set attribute clout, the effects are mirrored by its descendants. The leaf nodes compose of the end point classes. The best test condition or decision must be granted for every node in the sequence. For a given node t, where p(j|t) is the conditional probability of the class j.

$$GINI(t) = 1 - \sum_j [p(j|t)]^2 \quad (7)$$

- **XGBoost**: Extreme Gradient Boosting is a collaborative library in supporting distributed vectors generating a predictive model, which is a congregation of low decision-making bodies for estimation. It is very easy to produce predictions and fairly simple to practice in comparison with other learning algorithms. This can manage tiny volumes of data, which have problems in other algorithms as well as in a wide range of circumstances. We use the K model ensemble by incorporating the following outputs

$$\hat{y}_i = \sum_{K=1}^{K} f_k(x_i),\ f_k \in F \quad (8)$$

where F is the space of trees, x_i is the input, ŷi is the final output. We attempt to minimize the following loss function

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

$$L(\Phi) = \sum_i l(\hat{y}_i, y_i) + \sum_i \Omega(f_k) \quad (9)$$

where $\Omega(f) = \Upsilon T + \frac{1}{2}\lambda \|w\|^2$

Υ and λ are the hyper parameter

T is number of tree node

W is the vector of nodes

- **Support Vector Machines:** A SVM method is essentially an interpretation in the multidimensional space of various classes on a hyper-plane. In order to minimize the miscalculation, SVM must iteratively generate the hyperplane. SVM uses a technique for kernel tricks in order to turn the kernel into a larger dimension input space. The hyper-plane equation is as follows

$$w^T x + b = 0 \quad (10)$$

Where w is the weight vector and b is the bias

We are trying to optimize the distance between the data points and the hyper-plane. The loss function which helps to optimize the margin is the loss of the hinge.

$$L(w) = \sum_{i=1} \max(0, 1 - y_i[w^T x_i + b]) + \lambda\|w\|_2^2 \quad (11)$$

where $\max(0, 1 - y_i[w^T x_i + b])$is a Loss Function and $\max + \lambda\|w\|_2^2$ is a Regularization Factor.

IV. Experimental Evaluation

Evaluations are discussed in this segment of all algorithms for both datasets. In Table 2 we conclude that with the simplified Twitter dataset-1 we obtained the maximum accuracy with 94.91% and that with the measurement criterion selected for the accuracy with BOW functions of the SVM classification the same classification provided 73.12%. Table 1 with a bold typeset displays the strongest classification classifier score. It was also observed that linear regression and MNB classifier for IMDb film Dataset-2 provide stronger estimates of precision and accuracy respectively at 84.18% and 80.37%. The LR classifier has reached the maximum degree of precision by 94.99% and more than 94.02% with respect to table 3 and TF-IDF as functions. SVM forecasts the highest grades when using precision as an assessment criteria. Table 1 with a bold form range indicates the better test classification outcome. It was also observed that SVM classifiers were superior to the other classifiers of Dataset-2 in IMDb movies, with higher scores of 84.13% and 78.76% for precision and accuracy, respectively. From the empirical evidence as seen from Tables 2 and 3, it has been found that the SVM classifier is the most suitable classifier for twitter sentiment analysis. Evaluation of tweets for content analysis using ML Models is as shown in Fig 2.

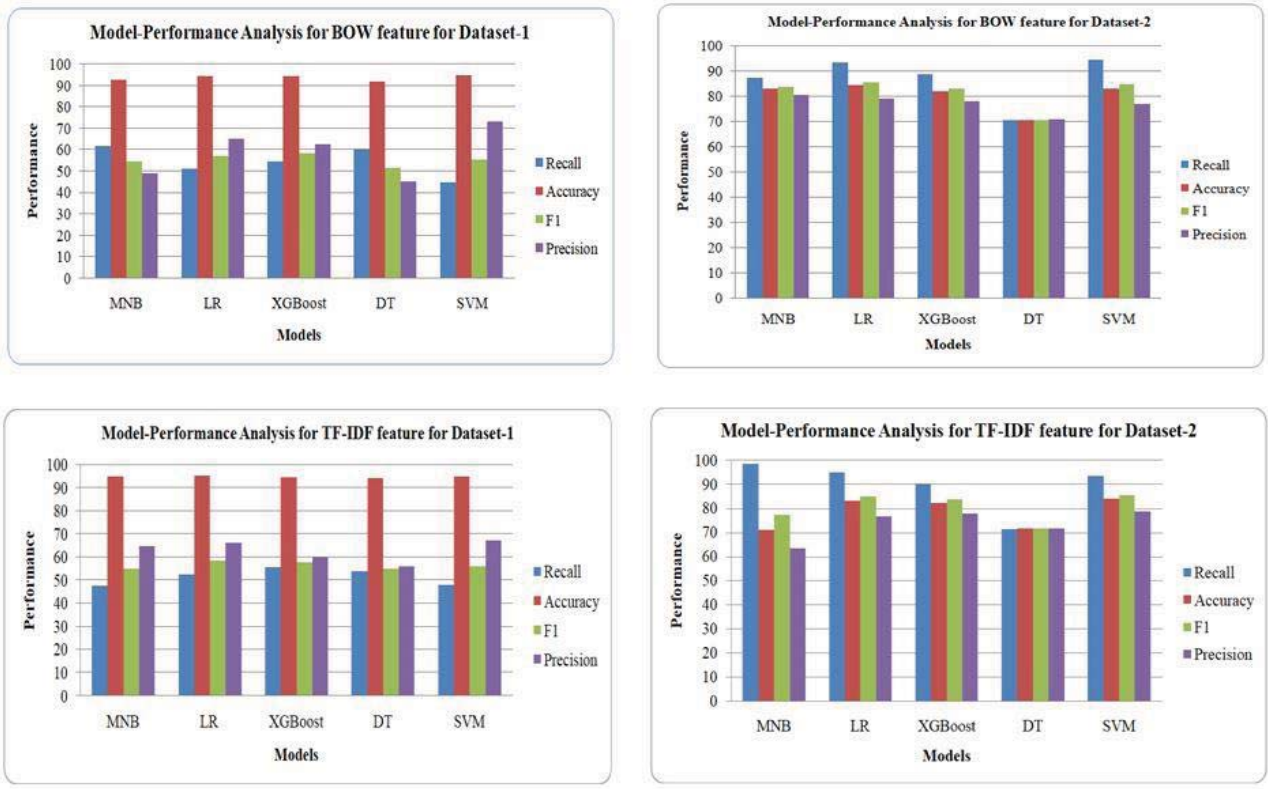


Fig 2: Performance of different models with respect to feature selected and dataset

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

TABLE 2: PERFORMANCE COMPARISON WITH DIFFERENT CLASSIFICATION METHODS ON THE DATASETS FOR BOW FEATURE

Parameter	MNB		LR		XGBoost		DT		SVM	
BOW	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2
Recall	**61.56**	87.24	51.10	93.21	54.34	88.46	60.08	70.30	44.47	**94.29**
Accuracy	92.75	82.93	94.58	**84.18**	94.45	81.74	91.95	70.48	**94.91**	82.75
F1	54.60	83.66	57.21	**85.51**	**58.11**	82.92	51.41	70.47	55.31	84.56
Precision	49.06	**80.37**	64.98	78.99	62.43	78.03	44.93	70.63	**73.12**	76.65

TABLE 3: PERFORMANCE COMPARISON WITH DIFFERENT CLASSIFICATION METHODS ON THE DATASETS FOR TF-IDF FEATURE

Parameter	MNB		LR		XGBoost		DT		SVM	
TF-IDF	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2	Dataset-1	Dataset-2
Recall	47.83	**98.59**	52.46	95.02	**55.86**	90.39	54.01	71.44	47.99	93.44
Accuracy	94.73	71.09	**94.99**	83.18	94.51	82.43	94.02	71.66	94.92	**84.13**
F1	55.11	77.32	**58.62**	84.95	57.92	83.73	54.98	71.59	56.08	**85.48**
Precision	64.98	63.60	66.40	76.82	60.13	77.97	56.00	71.75	**67.46**	**78.76**

V. CONCLUSION AND FUTURE SCOPE

The key objective of this research was to evaluate the ML classifier efficiency by choosing performance measurements. A differentiation is made from the results obtained between the model and the SVM model is described as best suited for Twitter knowledge classification. The proposed model effort went through the preprocessing level, the development of features and the classifiers' learning. As discussed earlier, the conceptual framework is analytically assessed in terms of evaluation parameters. It also explained that the output of the classifier not only depends on the dataset used, but also on the distribution of positive and negative data. Dataset-1 which has uneven labeled data as compared to the Dataset-2 out-performed on the various evaluation parameters.

As there are minimal features for learning the classifiers in our approach, we can use more numbers of features such as Information Gain, Chi-Square etc. There are several Pre-Processing techniques for dataset cleaning and we employ five major techniques. We can also utilize comprehensive learning approaches to obtain a stronger outcome for more sophisticated pre-processing methods not used in this analysis. In comparison to identification of emoji performance, the above study predicted text details more reliably. Taking into account emoji labeling and compiling a larger number of emoji datasets and proper labeling may help boost future research results. In addition, potential work to boost the quality using ensemble methods, other supervised, semi-supervised and reinforcement models.

.

REFERENCE

[1] Guo, and T. Yang, "Research and improvement of feature words weight based on TFIDF algorithm", IEEE Information Technology, Networking, Electronic and Automation Control Conference, Chongqing, pp. 415-419, 2016, doi: 10.1109/ITNEC.2016.7560393.

[2] Hasan, S. Moin, A. Karim, and S. Shamshirband, "Machine Learning-Based Sentiment Analysis for Twitter Accounts" Mathematical and Computational Applications, 23(1), 11, 2018, doi: 10.3390/mca23010011.

[3] Maryam, and R. Ali, "Temporal TF-IDF-Based Twitter Event Summarization Incorporating Keyword Importance", Smart Innovation, Systems and Technologies, pp 559–566, 2018, doi: 10.1007/978-981-13-1747-7_54.

[4] Das, and S. Chakraborty, "An Improved Text Sentiment Classification Model Using TF-IDF and Next Word Negation", 2018, arXiv: 1806.06407.

[5] Pang, L. Lee, "Opinion mining and sentiment analysis". Foundation Trends Information Retrieval 2(1–2), pp 1–135, 2008.

[6] Pujari, Aiswarya, and NP. Shetty, "Comparison of Classification Techniques for Feature Oriented Sentiment Analysis of Product Review Data", Data Engineering and Intelligent Computing, pp. 149–158, 2017, doi: 10.1007/978-981-10-3223-3_14.

[7] CN. Kamath, SS. Bukhari, and A. Dengel, "Comparative Study between Traditional Machine Learning and Deep Learning Approaches for Text Classification", Proceedings of the ACM Symposium on Document Engineering- DocEng, 2018, doi: 10.1145/3209280.3209526.

[8] Effrosynidis, S. Symeonidis, and A. Arampatzis, "A Comparison of Pre-processing Techniques for Twitter Sentiment Analysis", Lecture Notes in Computer Science, 394–406, 2017, doi: 10.1007/978-3-319-67008-9_31.

[9] D. Chatzakou, and A. Vakali, "Harvesting Opinions and Emotions from Social Media Textual Resources", IEEE INTERNET COMPUTING pp 46-50, 2015.

[10] Nazir, MA. Ghazanfar, M. Maqsood, F. Aadil, S. Rho, and I. Mehmood, "Social media signal detection using tweets volume, hashtag, and sentiment analysis", Multimedia Tools and Applications, 2018, doi: 10.1007/s11042-018-6437-z.

[11] Yamout, and R. Lakkis, "Improved TFIDF weighting techniques in document Retrieval", Thirteenth International Conference on Digital Information Management (ICDIM), 2018, doi: 10.1109/icdim.2018.8847156.

[12] Sidorov, "Vector Space Model for Texts and the tf-idf Measure", In: Syntactic n-grams in Computational Linguistics. Springer Briefs in Computer Science, pp 11–15, 2019, doi: 10.1007/978-3-030-14771-6_3.

[13] GM. Emelyanov, DV. Mikhailov, and A.P. Kozlov, "The TF-IDF measure and analysis of links between words within N-grams in the formation of knowledge units for open tests", Pattern Recognition and Image Analysis. 27, 825–831, 2017, doi: 10.1134/S1054661817040058.

[14] Ismail, S. Harous, and B. Belkhouche, "A Comparative Analysis of Machine Learning Classifiers for Twitter Sentiment Analysis", International Conference on Intelligent Text Processing and Computational Linguistics – CICLing, 2016.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

[15] HD. White, “Bag of works retrieval: TF*IDF weighting of works co-cited with a seed”, International Journal of Digital Library 19, pp 139–149, 2018, doi: 10.1007/s00799-017-0217-7.
[16] Alsmadi, and GK. Hoon, Term weighting scheme for short-text classification: Twitter corpuses. Neural Computing and Applications. 2018, doi: 10.1007/s00521-017-3298-8.
[17] Chen, C. Chen and Y. Liang, “Optimized TF-IDF Algorithm with the Adaptive Weight of Position of Word”, Advances in Intelligent Systems Research, volume 13, pp 114-117, 2016, doi: 10.2991/aiie-16.2016.28.
[18] J. Ho, D. Ondusko, B. Roy, and DF. Hsu, “Sentiment Analysis on Tweets Using Machine Learning and Combinatorial Fusion”, IEEE Intl Conf on Dependable, Autonomic and Secure Computing, 2019, doi: 10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00191.
[19] Dhanjal, and Sangeeta, “Applying Machine Learning Algorithms for News Articles Categorization: Using SVM and kNN with TF-IDF Approach”, Smart Computational Strategies: Theoretical and Practical Aspects, pp 95–105, 2019, doi: 10.1007/978-981-13-6295-8_9.
[20] Bilgin, and H. Köktaş, “Sentiment Analysis with Term Weighting and Word Vectors”, The International Arab Journal of Information Technology, Vol. 16, No. 5, pp 953-959, 2019.
[21] Nivaashini, RS. Soundariya, and P. Thangaraj, “Comparative Analysis of Machine Learning Approaches for Twitter Sentiment Analysis”, Journal of Computational and Theoretical Nanoscience, 15(5), pp 1743–1749, 2018, doi: 10.1166/jctn.2018.7371.
[22] M. Hasan, M Maliha, and M. Arifuzzaman, “Sentiment Analysis with NLP on Twitter Data”, International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), 2019, doi: 10.1109/ic4me247184.2019.9036670.
[23] Hassan, W. Gomaa, G. Khoriba, and M. Haggag, “Credibility Detection in Twitter Using Word N-gram Analysis and Supervised Machine Learning Techniques”, International Journal of Intelligent Engineering and Systems, Vol.13, No.1, 2020, doi: 10.22266/ijies2020.0229.27.
[24] R. Othman, Y. Abdelsadek, K. Chelghoum, I. Kacem, and R. Faiz, “Improving Sentiment Analysis in Twitter Using Sentiment Specific Word Embeddings”, 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2019, doi: 10.1109/idaacs.2019.8924403.
[25] S. El Mrabti, M. Al Achhab, and M. Lazaar, “Comparison of Feature Selection Methods for Sentiment Analysis”, Big Data, Cloud and Applications, pp 261–272, 2018, doi: 10.1007/978-3-319-96292-4_21.
[26] S. Mestry, H. Singh, R. Chauhan, V. Bisht, and K. Tiwari, “Automation in Social Networking Comments With the Help of Robust fastText and CNN”, 1st International Conference on Innovations in Information and Communication Technology, 2019, doi: 10.1109/iciict1.2019.8741503.
[27] SW. Kim, and JM. Gil, “Research paper classification systems based on TF-IDF and LDA schemes”, Human-Centric Computing and Information Sciences, 9(1), 2019, doi: 10.1186/s13673-019-0192-7.
[28] T. Renault, “Sentiment analysis and machine learning in finance: a comparison of methods and models on one million messages”. Digital Finance, 2019, doi: 10.1007/s42521-019-00014-x.
[29] Kshirsagar, “detecting hate tweets-twitter sentiment analysis”, https://towardsdatascience.com/detecting-hate-tweets-twitter-sentiment-analysis-780d8a82d4f6, (online last accessed on 06/06/2020)
[30] A. Pimpalkar, et al, “Political Comments Analysis Using Text Mining”, Recent Trends in Computer Science and Software Technology, Mantech Publications, Volume 3 Issue 1, pp 1-8, 2018.
[31] P. Shende, A. Pimpalkar, “Opinion Analysis of Public Sentiment on Social Websites using Web Mining and Natural Language Processing”, International Journal for Research in Emerging Science and Technology (IJREST), Special-Issue, pp 136-143, 2017.
[32] W. Gao, M. Peng, H. Wang, Y. Zhang, Q. Xie and G. Tian, “Incorporating word embeddings into topic modeling of short text”, Knowledge and Information Systems 61, 1123–1145, 2019, doi: 10.1007/s10115-018-1314-7.
[33] A. Pimpalkar, T. Wandhe, M. Swati and M. Kene, “Review of Online Product using Rule Based and Fuzzy Logic with Smileys”, International Journal of Computing and Technology (IJCAT), Vol. 1, Issue 1, pp 39-44, 2014.
[34] D. Chafale, A. Pimpalkar, “Sentiment Analysis on Product Reviews Using Plutchik’s Wheel of Emotions with Fuzzy Logic”, International Journal of Engineering & Technology (AIJET), Vol. 1 Issue 2, pp 1-8, 2014.
[35] R. Bawankule, A. Pimpalkar, “Sentence Level Text Extraction using Hierarchical Fuzzy Relational Clustering Algorithm”, International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 11, pp 91-95, 2014.
[36] A. Pimpalkar, “A Sentimental Analysis of Movie Reviews Involving Fuzzy Rule-Based”, International Journal of Artificial Intelligence and Knowledge Discovery (IJAIKD), Vol. 3, No. 2, pp 9-14, 2013.
[37] Y. Gu, Y. Wang, J. Huan, Y. Sun, and W. Jia, “ An Improved TFIDF Algorithm Based on Dual Parallel Adaptive Computing Model”, IEEE International Conference on Internet of Things, 2018, doi: 10.1109/cybermatics_2018.2018.00133.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:09:49 UTC from IEEE Xplore. Restrictions apply.

Microbiological
Communication

BBRC
Bioscience Biotechnology
Research Communications

Biosc.Biotech.Res.Comm. Vol 13 (3) July-Aug-Sep 2020 Pp-1156-1161

Evaluation of Inhibitory Activity of Bacteriocins from *Enterococcus italicus* BLN48 Against *Mycobacterium fortuitum* and its Toxicity Profiling

Revathy Kalyanasundaram, Manikkam Radhakrishnan* and Sivaraj Anbarasu
[1]*Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India*

ABSTRACT

Lactic Acid Bacteria (LAB) is an important group of microorganism due to their wide application in the food and dairy industries. They have extensively studied for the potential against various bacterial pathogens. The inhibitory activity of LAB is through their production of bacteriocins, organic acids, enzymes, hydrogen peroxide, etc., Bacteriocins are ribosomally synthesized antimicrobial peptides produced by various microorganisms. The incidence of non tuberculous mycobacterial infections increasing worldwide. *Mycobacterium fortuitum* is one of the rapidly growing non tuberculous mycbacteria which causes skin, bone, joint and pulmonary infections. Their antibiotic resistance and prolonged course of treatment necessitates the development of new candidate to fight against them. In this study, we have partially purified bacteriocins from four LAB strains using solvent extraction method and screened their activity against *M. fortuitum* MTCC1902 by colony forming unit (CFU) estimation method. The potential strain was identified by 16S rRNA sequencing and their sequences were submitted to Genbank database. The toxicity of potential LAB strain was assessed by both in vitro and in vivo method against Vero cell lines and zebra fish model respectively.The strain BLN48 showed 97.9% reduction in growth of *M. fortuitum* and 1.89±0.32 log reduction in CFU/ml. The potential strain BLN48 was identified as *Enterococcus italicus*. *E. italicus* BLN48 exhibited cytotoxicity against vero cell lines in dose dependent manner whereas under in vivo conditions, 50% of zebrafish larvae survived upto 144 hours post fertilization (hpf) with normal morphological changes. Further purification and characterization of *E. italicus* BLN48 bacteriocin in future helps in the development of an efficient candidate against *M. fortuitum* as well as other mycobacterial pathogens.

KEY WORDS: PARTIALLY PURIFIED BACTERIOCIN, NON TUBERCULOUS MYCOBACTERIA, MYCOBACTERIUM FORTUITUM, CYTOTOXICITY, ZEBRAFISH.

Article Information:
Corresponding author email: *mrkactinos@gmail.com*
Received 11/07/2020 Accepted after revision 03/09/2020
P-ISSN: 0974-6455 E-ISSN: 2321-4007
Thomson Reuters ISI Clarivate Analytics
Web of Science ESCI Indexed Journal

Identifiers & Pagination:
Vol 13(3) E-Pub 30th Sep 2020 Pp- 1156-1161
This is an open access article under Creative Commons License Attribution International (CC-BY 4.0)
Published by Society for Science & Nature India
DOI: *http://dx.doi.org/10.21786/bbrc/13.3/25*

INTRODUCTION

Lactic Acid Bacteria (LAB) remains an industrially important group of microorganism due to their wide application in the food and dairy industries (Gomez et al., 2015). LAB are widely distributed in diverse habitats like marine, food products, dairy origin, etc., Usually, fermented foods were screened for bacterial isolates with antimicrobial properties as its microbiota is dominated by LAB. LAB also exists in gastrointestinal tracts, oral cavities of humans as well as animals (Bungenstock et al., 2020; Li et al., 2020). In raw milk, LAB are the predominant microorganisms and they contribute to the fermentation and food preservation process due to their various metabolite production (Rahmeh et al., 2019). LAB has studied extensively for their antagonistic activity against various bacterial pathogens such as food spoiling microorganisms like *Listeria monocytogenes,* gastrointestinal pathogens and other various gram positive and gram negative bacterial pathogens. Researchers have gained significant attention towards LAB due to their Generally Recognized As Safe (GRAS) status. The inhibitory activity of LAB is mainly through their production of various substances like organic acids, enzymes, bacteriocins, hydrogen peroxide, etc., (Gupta and Garg 2009; Rodrigues et al., 2006 Hussein et al., 2018).

;Bacteriocins are ribosomally synthesized antimicrobial peptides produced by various microorganisms. The uses of bacteriocins in functional foods and as an alternative to antibiotics are their emerging application (Quwehand et al., 2004; Messi et al., 2001). Physical stability and non-toxic nature are the major advantages of bacteriocins (Morgan et al., 2005). Different classes of bacteriocins exerts diverse mechanisms against their target like disruption of cell wall, pore formation, inhibition of protein and nucleic acid synthesis, etc., (Cascales et al., 2007; Stevens et al., 1991). LAB produces a diverse nature of bacteriocins in different size, physicochemical properties, spectrum of activity, structures, etc. They secreted in an extracellular space during LAB growth (Venegas et al., 2019; Anbarasu et al., 2020). The crude and purified bacteriocins from LAB have found their potential use as biopreservative agents to enhance the quality and safety of various food products.The potential bacteriocins can be used in both combined and balanced mode as probiotics for human diseases also (Iseppi et al., 2019, Arrioja et al., 2020).

Non Tuberculous Mycobacteria (NTM) are several mycobacterial species other than *M. tuberculosis* complex and *M. leprae*. NTM causes opportunistic infections in humans as well as animals and also it transmitted among environment, livestock, wildlife, etc. (Odoi et al., 2020). *Mycobacterium fortuitum* is one of the rapid growing NTM and is predominantly found in water systems like natural water, tap water, and water used in showers in hospitals and soil. It mainly causes skin, bone, joint infections and pulmonary diseases in immunocompromised and immunosuppressed patients. They also causes surgical site infections (Okamori et al., 2018; Griffith et al., 2007; Goslee and Wolinsky 1976; Wolinsky and Rynearson 1968; Choudhary et al., 2020).

M. fortuitum are often isolated from skin and soft tissues and also from other clinical samples as it causes many types of infection (Garcia et al., 2020). The development of antibiotic resistance and prolonged course of treatment with multiple antibiotics in NTM infection surges the need for the development of new candidates with potential inhibitory substance and less toxicity to fight against these infections. In this study, we have evaluated bacteriocins from four LAB isolates for their anti *M. fortuitum* activity. The potential LAB isolate was identified through 16S rRNA sequencing and phylogenetic analysis. Toxicity profile of the bacteriocins from the potential LAB strain was also evaluated through in vitro and in vivo method.

MATERIAL AND METHODS

Mycobacterium fortuitum MTCC1902 strain was purchased from Microbial Type Culture Collection (MTCC), Chandigarh, India. de Man Rogosa Sharpe broth (Himedia), Middlebrook 7H9 broth (Himedia), Middlebrook 7H11 Agar (Himedia), Chloroform (SRL), PBS tablets (Sigma) were used in the study. Four lactic acid bacterial strains viz., BLN 34, BLN 36, BLN 39 and BLN 48 previously isolated from different cow milk samples were used in this study. Viability of all the cultures were maintained in MRS agar slants at 4°C (Revathy et al., 2019). Bacteriocin from four selected cultures was produced by submerged fermentation process and was partially purified by solvent extraction method using chloroform as described by Burianek et al., 2000 with few modifications. Briefly, 5ml of overnight grown culture in MRS broth was added to 500ml of sterile MRS broth and incubated for 18 hours in shaking incubator at 30˚C. After incubation, the culture was centrifuged at 5000rpm for 10minutes to collect the cell free supernatant. For the extraction of crude bacteriocin, 50% v/v of chloroform has been added to the supernatant and kept in magnetic stirrer at 1000rpm for 20minutes. Then the mixture was subjected to centrifugation at 10000rpm for 30 minutes. After centrifugation, the precipitate in the interphase layer between solvent and aqueous phase was collected carefully and freezed at -20˚C following by lyophilization. The lyophilized form of partially purified bacteriocin (PPB) were stored at -20˚C and used for further assays.

Evaluation of PPB against *M. fortuitum:* Ten mg concentration of partially purified bacteriocin in the form of lyophilized powder was dissolved in 1ml of PBS buffer in order to get 10mg/ml (w/v)concentration. Desired working concentration of PPB was prepared from the main stock using PBS buffer. Inhibitory activity of PPB prepared from all the four cultures was evaluated against *M. fortuitum* by colony forming unit (CFU) estimation (Gillespie et al. 2005).Briefly, *M. fortuitum* MTCC1902 suspension was prepared by inoculating a loopful of *M. fortuitum* culture into 0.3ml of Middlebrook 7H9

broth followed by vortexing. Then the volume of the suspension was made upto 5ml. using 7H9 broth. In a sterile cryovial, 400µl of 7H9 broth was used as growth control and 350µl of 7H9 broth with 50µl of bacteriocin was used in test vial. All the vials were added with 100µl of *M. fortuitum* suspension and incubated at 37˚C for 48 hours. After incubation, 100µl of aliquot from each vial was serially diluted in 900µl of PBS buffer upto 10-12 dilution. 50µl of the dilution was spreaded onto Middlebrook 7H11 agar plate. Plating was done in triplicate and all the plates were incubated at 37ºC for 48 hours. The plates with individual colonies were taken for the calculation of CFU/ml.

Taxonomy of potential LAB strain: The genomic DNA of LAB strain BLN48 was isolated using solute ready genomic DNA kit. DNA was analyzed by gel electrophoresis and quantified using a spectrophotometer (NanoDrop ND-1000, Thermo Scientific, Gloucester, UK). The 16S rRNA gene sequence of the strain was amplified using the primers: 27F 5´AGAGTTTGATCMTGGCTCAG3´ (forward) and 1492R 5´TACGGYTACCTTGTTACGACTT3´ (reverse) (Kumar Gothwal et al., 2007). The PCR amplified product of the strain was sequenced and analyzed at National Collection of Industrial Microorganisms (NCIM), National Chemical Laboratory (CSIR-NCL), Pune, India. The 16S rRNA gene sequence obtained from the strain BLN48 was aligned with similar sequences available in GenBank using MEGA 7 program. The aligned sequences of the strain BLN48 was used to construct the phylogenetic tree by following neighbor joining algorithm in MEGA 7 program (Saitou and Nei, 1987). The bootstrap estimation (Felsenstein, 1985) was used to determine the confidence of the branches of the phylogenetic tree. The partial 16S rRNA nucleotide sequence of all the four strains has been deposited in GenBank database.

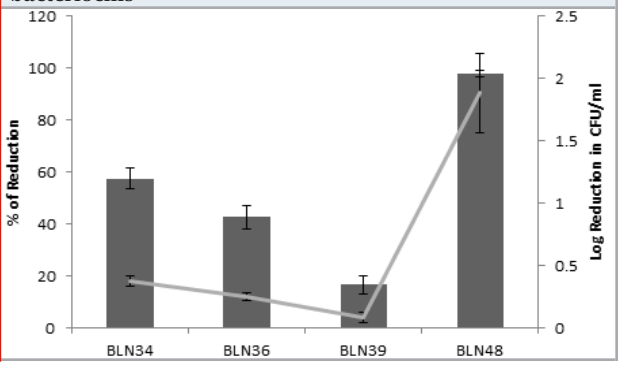
In vitro toxicity analysis of PPB: The cytotoxicity of potential PPB which showed activity against *M. fortuitumwere* assessed in vitro by adopting MTT (3-(4, 5-dimethyl thiazol-2yl)-2, 5-diphenyl tetrazolium bromide) assay using Vero cell lines. Briefly, in the 96 well plate, 100µl of RMPI 1640 medium was added with 100µl of desired concentration of PPB. Then 200µl of total volume is gently mixed well. 100µl of diluted PPB from the first well was serially diluted in next well till reaching the lowest concentration. The cultured Vero cell lines were harvested by trypsinization and pooled in 50ml vial. Then the cells were plated at a density of 1x100cells/ml. 200µl of vero cells without PPB was used as a control. The cells were incubated at 37˚C in 5% CO2 incubator for 24 hours. After incubation, 20µl of MTT solution was added to all the wells and incubated for 4 hours at 37˚C.The media and MTT was well mixed and the absorbance was measured at 450nm and the percentage of viability was calculated manually (Vijayarathna and Sasidharan, 2012).

In vivo toxicity analysis of PPB: In vivo toxicity of PPB was evaluated using zebrafish as a model (Sisman et al. 2008). Zero day old zebrafish eggs were purchased from zebrafish aquarium in Kanchipuram district, Tamil Nadu, India. Twenty healthy post hatched zebra fish eggs were transferred to the wells of a 24-well plate along with 1 ml of embryo water (60 mg of sea salt/ litre of ultrapure water). Different concentrations of PPB of potential LAB strain BLN48 (10, 50 and100 µg/ml) was added to the wells and incubated for 144 h at 28.5ºC. Mortality of the zebra fish was noted after 24, 48, 96 and 144 h. The embryos appeared opaque and white in colour.The dead embryos were degraded soon, whereas the structures of intact embryos were more visible by 48 hours post fertilization (hpf) which allowed a clear distinction between the dead and alive.The mortality rate is calculated. At the end of the incubation period, the embryos were photographed using a light microscope at 10X magnification.

RESULTS AND DISCUSSION

Anti M. *fortuitum* **activity:** Among the four isolates tested, BLN48 showed significant inhibitory activity against *M. fortuitum* MTCC1902 viz., 1.89 log reduction of CFU/ml which corresponds to 97.9% reduction in growth when compared to growth control. Followed by, BLN34 showed slight inhibitory activity by 0.38 log reduction of CFU/ml with 57.56% reduction from growth control (Figure 1).

Figure 1: Anti *M. fortuitum* activity of Partially purified bacteriocins



Taxonomy of potential LAB strain BLN48: Amplification of 16S rRNA gene from the strain BLN48 resulted in 1435 bp sequences. BLAST analysis showed 99.43% sequence similarity with 16S rRNA gene sequence of *Enterococcus italicus* DSM 15952. The phylogenetic tree also showed that the strain BLN48 is closely related to *Enterococcus italicus* (Figure 2). The nucleotide sequence of *E. italicus* BLN48 was submitted to Genbank with accession number MN880432.

Table 1: In vitro cytotoxicity analysis of BLN48 by MTT assay against vero cell lines

PPB	Test concentration	Percentage of Viability
BLN48	100mM	16.6
	10mM	60.4

In vitro toxicity analysis of PPB: The cytotoxicity analysis by MTT assay showed that the bacteriocin from BLN48 exhibited the cytotoxicity on vero cell lines in dose dependent manner. At the maximum of 60% of vero cells survived when treated at 10mM concentration of bacteriocin (Table 1).

Figure 2: Phylogeny of BLN48



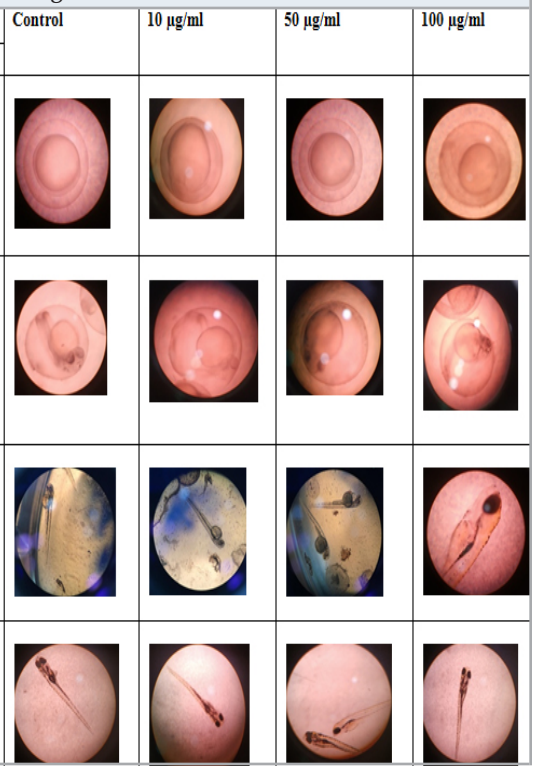
In vivo toxicity analysis of PPB: In the in vivo toxicity analysis of BLN48 done with zebrafish larvae, BLN48 have shown less toxicity viz., around 50% of larvae (compared to control) were survived upto 144 hpf at high concentration (100µg/ml) with healthy morphology under microscopic observation (Table 2). The features like fin movement, swimming nature, tail development are normal to the viable larvae (Figure 3).

Table 2. Mortality rate in in vivo toxicity analysis of BLN48

Concentration HPF	Control	10 µg/ml	50 µg/ml	100 µg/ml
0 HPF	20	20	20	20
24 HPF	20	18	19	16
48 HPF	18	17	15	13
96 HPF	16	14	12	9
144 HPF	13	11	9	7

M. fortuitum group is responsible for 60-80% of post surgical infections caused by mycobacteria. The successful treatment outcome of *M. fortuitum* infections often limited by the multi drug resistance, need of combination therapy, prolonged course, etc. (Cynamon et al., 2012).Santos et al (2016) analysed the resistant profile of *M. fortuitum* isolates and showed their resistance to different classes of antibiotics. Antimicrobial peptides i.e., bacteriocins offers a solution to combat antibiotic resistance of various pathogenic microorganisms. Numerous studies have proved the efficacy of bacteriocins against various drug resistant pathogens (Regmi et al., 2017).

Figure 3: Morphological changes in zebrafish embryos after treating with BLN48



In the present study, the partially purified bacteriocin from *Enterococcus italicus* BLN48 has showed significant inhibition against *M. fortuitum*. In the various studies, numerous antibiotics and chemical compounds have been screened against *M. fortuitum* but there are very less reports on the screening of natural compounds of microbial source against *M. fortuitum* (Gay et al., 1984; Welch et al., 1979; Bagchi et al., 2007). While there are many other reports focuses on the screening of bacteriocins against other mycobacterium species like *M. tuberculosis*. A study by Sosunov et al (2007) has assessed the antimycobacterial of five bacteriocins against *M. tuberculosis* strains. In 2010, Carroll et al., compared the activities of two bacteriocins Lacticin 3147 and nisin against NTM species like *M. kansasii, M. avium* paratuberculosis.

Their study found that lacticin 3147 showed superior activity than nisin. A study by Aguilar-Pérez et al (2018) have found the inhibitory activity of bacteriocin AS-48 against *M. fortuitum* and found their MIC as 64µg/ml along with other mycobacterium species. They also found that there is no cytotoxicity obtained against various macrophage cell lines. Our result showed that the *E. italicus* BLN48 exhibits cytotoxicity on vero cell lines in dose dependent manner. However, the in vivo cytotoxicity assay using zebrafish showed that there is no side effect on the morphology of viable larvae in the presence of *E. italicus* BLN48 at three different concentrations. A survey by Fortina et al (2008) has described the safety and biotechnological properties of *E. italicus* of dairy origin. They suggested that *E.*

italicus presence in the cheese lowers the health risk and supports their applications in dairy industry. This study describes the safety profile of *E. italicus* for their wide application in future.

CONCLUSION

Mycobacterium fortuitum is one of the clinically significant rapidly growing mycobacteria which cause pulmonary, skin and soft tissue infections in immunocompromised and immunosuppressed patients. Partially purified bacteriocin of *Enterococcus italicus* BLN48 showed significant inhibitory activity against *M. fortuitum*. Their cytotoxicity assay under both in vitro and in vivo conditions shows that *E. italicus* BLN48 can be developed as potential candidate against *M. fortuitum*. Further analysis on their Minimum Inhibitory Concentration (MIC), purification and their characterization and screening them against other mycobacterial species will leads to the efficient use of *E. italicus* BLN48 for therapeutic purpose.

REFERENCES

Aguilar-Pérez, C., Gracia, B., Rodrigues, L., Vitoria, A., Cebrián, R., Deboosère, N., Song, O.R., Brodin, P., Maqueda, M. & Aínsa, J.A. (2018) Synergy between circular bacteriocin AS-48 and ethambutol against *Mycobacterium tuberculosis*. Antimicrobial agents and chemotherapy, 62(9).

Anbarasu, S., Kumar, V., Rani, U., Kalyanasundram, R. & Radhakrishnan, M. (2020) Purification and characterization of bacteriocins active against *Mycobacterium tuberculosis* from lactic acid bacteria. Research Journal of Biotechnology, 15, 4.

Arrioja, D., Mani-López, E., Palou, E., & López-Malo, A. (2020). Antimicrobial activity and storage stability of cell-free supernatants from lactic acid bacteria and their applications with fresh beef. Food Control, 107286.

Bagchi, M. C., Mills, D., & Basak, S. C. (2007). Quantitative structure-activity relationship (QSAR) studies of quinolone antibacterials against *M. fortuitum* and *M. smegmatis* using theoretical molecular descriptors. Journal of molecular modeling, 13(1), 111-120.

Bungenstock, L., Abdulmawjood, A., & Reich, F. (2020). Evaluation of antibacterial properties of lactic acid bacteria from traditionally and industrially produced fermented sausages from Germany. Plos one, 15(3), e0230345.

Burianek, L. L., & Yousef, A. E. (2000). Solvent extraction of bacteriocins from liquid cultures. Letters in applied microbiology, 31(3), 193-197.

Carroll, J., & O'Mahony, J. (2011). Anti-mycobacterial peptides: made to order with delivery included. Bioengineered bugs, 2(5), 241-246.

Carroll, J., Draper, L.A., O'Connor, P.M., Coffey, A., Hill, C., Ross, R.P., Cotter, P.D. and O'Mahony, J. (2010) Comparison of the activities of the lantibiotics nisin and lacticin 3147 against clinically significant mycobacteria. International journal of antimicrobial agents, 36(2), 132-136.

Cascales, E., Buchanan, S.K., Duché, D., Kleanthous, C., Lloubes, R., Postle, K., Riley, M., Slatin, S. & Cavard, D. (2007) Colicin biology. Microbiology and molecular biology reviews, 71(1), 158-229.

Choudhary, A., Gopalakrishnan, R., Senthur Nambi, P., Thirunarayan, M.A., Ramasubramanian, V. & Sridharan, S. (2020) Surgical Site Infections Caused by Rapidly Growing Nontuberculous Mycobacteria: an Under-Recognized and Misdiagnosed Entity. Indian Journal of Surgery, 1-6.

Cynamon, M., Jureller, J., Desai, B., Ramachandran, K., Sklaney, M., & Grossman, T. H. (2012). In vitro activity of TP-271 against *Mycobacterium abscessus*, *Mycobacterium fortuitum*, and *Nocardia* species. Antimicrobial agents and chemotherapy, 56(7), 3986-3988.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 39(4), 783-791.

Fortina, M.G., Ricci, G., Borgo, F., Manachini, P.L., Arends, K., Schiwon, K., Abajy, M.Y. & Grohmann, E. (2008) A survey on biotechnological potential and safety of the novel *Enterococcus* species of dairy origin, *E. italicus*. International journal of food microbiology, 123(3), 204-211.

Garcia-Coca, M., Rodríguez-Sevilla, G., Pérez-Domingo, A., Aguilera-Correa, J.J., Esteban, J. & Muñoz-Egea, M.C. (2020) Inhibition of *Mycobacterium abscessus, M. chelonae*, and *M. fortuitum* biofilms by *Methylobacterium* sp. The Journal of Antibiotics, 73(1), 40-47.

Gay, J. D., DeYoung, D. R., & Roberts, G. D. (1984). In vitro activities of norfloxacin and ciprofloxacin against *Mycobacterium tuberculosis*, *M. avium* complex, *M. chelonei, M. fortuitum*, and *M. kansasii*. Antimicrobial agents and chemotherapy, 26(1), 94-96.

Gillespie, S. H., Basu, S., Dickens, A. L., O'Sullivan, D. M., & McHugh, T. D. (2005). Effect of subinhibitory concentrations of ciprofloxacin on *Mycobacterium fortuitum* mutation rates. Journal of Antimicrobial Chemotherapy, 56(2), 344-348.

Gómez-Sala, B., Muñoz-Atienza, E., Sánchez, J., Basanta, A., Herranz, C., Hernández, P. E., & Cintas, L. M. (2015). Bacteriocin production by lactic acid bacteria isolated from fish, seafood and fish products. European Food Research and Technology, 241(3), 341-356.

Goslee, S., & Wolinsky, E. (1976). Water as a source of potentially pathogenic mycobacteria. American Review of Respiratory Disease, 113(3), 287-292.

Griffith, D.E., Aksamit, T., Brown-Elliott, B.A., Catanzaro, A., Daley, C., Gordin, F., Holland, S.M., Horsburgh, R., Huitt, G., Iademarco, M.F. & Iseman, M. (2007) An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. American journal of respiratory and critical care

medicine, 175(4), 367-416.
Gupta, V., & Garg, R. (2009). Probiotics. Indian journal of medical microbiology, 27(3), 202.
Hussein, A. R., Khalaf, Z. Z., Samir, Z., & Samir, R. (2018). Antibacterial activity of crud Bacteriocin-like substance against food borne bacterial pathogens. Iraqi journal of Science, 59(1A), 16-24.
Iseppi, R., Messi, P., Camellini, S. and Sabia, C. (2019) Bacteriocin activity of *Lactobacillus brevis* and *Lactobacillus paracasei ssp. paracasei*. Journal of medical microbiology, 68(9), 1359-1366.
Kumar Gothwal, R., Kumar Nigam, V., Mohan, M. K., Sasmal, D., & Ghosh, P. (2007). Extraction of bulk DNA from Thar Desert soils for optimization of PCR-DGGE based microbial community analysis. Electronic Journal of Biotechnology, 10(3), 400-408.
Li, M., Wang, Y., Cui, H., Li, Y., Sun, Y., & Qiu, H. J. (2020). Characterization of Lactic Acid Bacteria Isolated From the Gastrointestinal Tract of a Wild Boar as Potential Probiotics. Frontiers in veterinary science, 7, 49.
Messi, P., Bondi, M., Sabia, C., Battini, R., & Manicardi, G. (2001). Detection and preliminary characterization of a bacteriocin (plantaricin 35d) produced by a *Lactobacillus plantarum* strain. International Journal of Food Microbiology, 64(1-2), 193-198.
Morgan, S. M., O'Connor, P. M., Cotter, P. D., Ross, R. P., & Hill, C. (2005). Sequential actions of the two component peptides of the lantibiotic lacticin 3147 explain its antimicrobial activity at nanomolar concentrations. Antimicrobial agents and chemotherapy, 49(7), 2606-2611.
Odoi, J.O., Ohya, K., Moribe, J., Takashima, Y., Sawai, K., Taguchi, K., Fukushi, H., Wada, T., Yoshida, S. & Asai, T. (2020) Isolation and antimicrobial susceptibilities of Nontuberculous Mycobacteria from Wildlife in Japan. Journal of Wildlife Diseases.
Okamori, S., Asakura, T., Nishimura, T., Tamizu, E., Ishii, M., Yoshida, M., Fukano, H., Hayashi, Y., Fujita, M., Hoshino, Y. & Betsuyaku, T. (2018) Natural history of *Mycobacterium fortuitum* pulmonary infection presenting with migratory infiltrates: a case report with microbiological analysis. BMC Infectious Diseases, 18(1), 1.
Quwehand, A.C. & Vesterlund, S. 2004. Antimicrobial components from lactic acid bacteria. Lactic Acid Bacteria Microbiological and Functional Aspects, Marcel Dekker Inc., New York, USA.
Rahmeh, R., Akbar, A., Kishk, M., Al-Onaizi, T., Al-Azmi, A., Al-Shatti, A., Shajan, A., Al-Mutairi, S. & Akbar, B. (2019) Distribution and antimicrobial activity of lactic acid bacteria from raw camel milk. New microbes and new infections, 30, 100560.
Regmi, S., Choi, Y. H., Choi, Y. S., Kim, M. R., & Yoo, J. C. (2017). Antimicrobial peptide isolated from *Bacillus amyloliquefaciens* K14 revitalizes its use in combinatorial drug therapy. Folia microbiologica, 62(2), 127-138.
Revathy, K., Manikkam, R., & Anbarasu, S. (2019). Isolation, characterization of lactic acid bacteria from cow and buffalo milk and evaluation for antibacterial and antimycobacterial activity in vitro. Asian Journal of Microbiology Biotechnology and Environmental Sciences, 21(4), 1041-1046.
Rodrigues, L., Banat, I. M., Teixeira, J., & Oliveira, R. (2006). Biosurfactants: potential applications in medicine. Journal of Antimicrobial Chemotherapy, 57(4), 609-618.
Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution, 4(4), 406-425.
Santos, D. R. D. S., Lourenço, M. C. S., Coelho, F. S., Mello, F. C. Q., & Duarte, R. S. (2016). Resistance profile of strains of *Mycobacterium fortuitum* isolated from clinical specimens. Jornal Brasileiro de Pneumologia, 42(4), 299-301.
Sisman, T., Incekara, Ü., & Yıldız, Y. S. (2008). Determination of acute and early life stage toxicity of fat-plant effluent using zebrafish (*Danio rerio*). Environmental Toxicology: An International Journal, 23(4), 480-486.
Sosunov, V., Mischenko, V., Eruslanov, B., Svetoch, E., Shakina, Y., Stern, N., Majorov, K., Sorokoumova, G., Selishcheva, A. & Apt, A. (2007) Antimycobacterial activity of bacteriocins and their complexes with liposomes. Journal of antimicrobial chemotherapy, 59(5), 919-925.
Stevens, K. A., Sheldon, B. W., Klapes, N. A., & Klaenhammer, T. R. (1991). Nisin treatment for inactivation of *Salmonella* species and other gram-negative bacteria. Appl. Environ. Microbiol., 57(12), 3613-3615.
Venegas-Ortega M.G., Flores-Gallegos A.C., Martínez-Hernández, J.L., Aguilar C.N. & Nevárez-Moorillón, G.V. (2019) Production of bioactive peptides from lactic acid bacteria: a sustainable approach for healthier foods. Comprehensive Reviews in Food Science and Food Safety, 18(4), 1039-1051.
Vijayarathna, S., & Sasidharan, S. (2012). Cytotoxicity of methanol extracts of *Elaeis guineensis* on MCF-7 and Vero cell lines. Asian Pacific Journal of Tropical Biomedicine, 2(10), 826-829.
Welch, D. F., & Kelly, M. T. (1979). Antimicrobial susceptibility testing of *Mycobacterium fortuitum* complex. Antimicrobial agents and chemotherapy, 15(6), 754-757.
Wolinsky, E., & Rynearson, T. K. (1968). Mycobacteria in soil and their relation to disease-associated strains. American Review of Respiratory Disease, 97(6P1), 1032-1037.

Pharmaceutical
Communication

BBRC
Bioscience Biotechnology
Research Communications

Biosc.Biotech.Res.Comm. Vol 13 (2) April-May-June 2020 Pp-401-409

Evaluation of Anti-HIV Activity of Selected Medicinal Plants: A Short Review

Raghavi R[1], Keren Deborah S[1], Jerrine Joseph[2]* and Wilson Aruni[2,3,4]

[1,2]Centre for Drug Discovery and Development, Sathyabama Institute for Science and Technology (Deemed to be University) Chennai 600119, India

[3]School of Medicine, Loma Linda University CA, USA.

[4]Musculoskeletal Disease Research Laboratory, US Department of Veteran Affairs, Loma Linda CA, USA.

ABSTRACT

Human immunodeficiency virus (HIV) causes the potentially life-threatening and chronic disease called acquired immune deficiency syndrome (AIDS). The main target of this viral disease is to suppress the immune system and make the body unresponsive to external stimuli. According to global health observatory data since epidemic, more than 78 million people were affected by HIV and 39 million people died globally. There were approximately 37.9 million people living with HIV at the end of 2018. Currently, antiretroviral therapy (ART) is available for the control of HIV but has serious associated side effects such as lipodystrophy. Because of the limitations, associated with ART, researchers throughout the world are trying to explore and develop more reliable and safe drugs from natural resources to manage HIV infection. A wide range of medicinal plants have been studied and have reported significant potential against HIV. Medicinal plants contain novel anti-HIV compounds. As it has been well reported that medicinal plants contain various types of phytochemical constituents including alkaloids, flavonoids, phenolic compounds, glycosides, tannins, and saponins, hence the medicinal plants could be potential sources of boosting immune responses, as well as halting the replication of HIV. A literature survey of medicinal plants from PubMed and plant literature database, was carried out to identify the plants with novel antiviral agents reported for the treatment of HIV/AIDS worldwide. Bioactive compounds from plants which play effective roles in the management of AIDS, which have been discussed in this review study. This could pave way for being taken up for active future in vitro and preclinical research studies to qualify as lead anti HIV molecules which is the need of the hour.

KEY WORDS: AIDS, ANTIRETROVIRAL THERAPY, PHYTOCONSTITUENTS, MEDICINAL PLANTS.

Article Information:*Corresponding Author: *jerrine.jj@gmail.com*
Received 15/04/2020 Accepted after revision 15/06/2020
Published: 30th June 2020 Pp-401-409
This is an open access article under Creative Commons License,.
Published by Society for Science & Nature, Bhopal India.
Available at: https://bbrc.in/
Article DOI: http://dx.doi.org/10.21786/bbrc/13.2/07

INTRODUCTION

HIV continues to be a major global public health issue, having claimed more than 32 million lives so far. However, with increasing access to effective HIV prevention, diagnosis, treatment and care, including for opportunistic infections, HIV infection has become a manageable chronic health condition, enabling people living with HIV to lead long and healthy lives. There were approximately 37.9 million people living with HIV at the end of 2018.As a result of concerted international efforts to respond to HIV, coverage of services has been steadily increasing. In 2018, 62% of adults and 54% of children living with HIV in low- and middle-income countries were receiving lifelong antiretroviral therapy (ART).(WHO 2019). HIV is a retrovirus that can integrate its DNA into the host genome. The virus enters the host cell and affects the immune system mainly T lymphocytes, monocytes, macrophages and dendritic cells (Salehi et al,.2018).

Its genetic material RNA is made up of nine genes which contain all the instructions to make new viruses. Three of these genes – gag, pol and env – provide the instructions to make proteins that will form new virus particles. The other six genes rev, nef, vif, vpr and vpu, provide code to make proteins that control the ability of HIV to infect a cell, produce new copies of virus or release viruses from infected cells. The HIV-1 binds to the chemokine receptor 5 or the CXC chemokine receptor 4 by interacting with the envelope proteins to gain entry to the host cell (Salehi et al,.2018). Therapies are now available to inhibit various stages of viral infection such as entry inhibitors, reverse transcriptase inhibitors, integrase strand transfer inhibitors and protease inhibitors.Presence of antibody to HIV proteins is well accepted as indicative of HIV infection. Sometimes certain clinical conditions may also result in the presence of false-positive HIV antibody. Serologic tests for HIV includes ELISA, Western blot and HIV p24 antigen assay.

Types and Symptoms

1. Primary infection (Acute HIV): Some people infected by HIV develop a flu-like illness within two to four weeks after the virus enters the body. This illness, known as primary (acute) HIV infection, may last for a few weeks. Possible signs and symptoms include fever, headache, muscle aches and joint pain, rash, sore throat and painful mouth sores, swollen lymph glands, mainly on the neck, diarrhoea, weight loss, cough, night sweats. As the infection progressively weakens the immune system, they can develop other signs and symptoms, such as swollen lymph nodes, weight loss, fever, diarrhoea and cough. Without treatment, they could also develop severe illnesses such as tuberculosis (TB), cryptococcal meningitis, severe bacterial infections, and cancers such as lymphomas and Kaposi's sarcoma (WHO 2019).

2. Clinical latent infection (Chronic HIV): In this stage of infection, HIV is still present in the body and in white blood cells. However, many people may not have any symptoms or infections during this time.

Treatment: Despite challenges, new global efforts have meant that the number of people receiving HIV treatment has increased dramatically in recent years, particularly in resource-poor countries. In 2018, 62% of all people living with HIV were accessing treatment. Of those, 53% were virally suppressed. This equates to 23.3 million people living with HIV receiving antiretroviral treatment (ART) in 2018 – up from 7.7 million in 2010. However, this level of treatment scale up is still not enough for the world to meet its global target of 30 million people on treatment by 2020 (WHO 2019). Significant progress has been made in the prevention of mother-to-child transmission of HIV (PMTCT). In 2018, 82% of all pregnant women living with HIV had access to treatment to prevent HIV transmission to their babies – an increase of more than 90% from 2010.

Antiretroviral Therapy: The combination of drugs used to treat HIV is called antiretroviral therapy antiretroviral therapy (ART). ART is recommended for all people living with HIV, regardless of how long they've had the virus or how healthy they are. More than two dozen antiretroviral drugs has been approved by FDA to treat HIV infection. Different classes of antiretroviral drugs act at different stages of the HIV life cycle. Two nucleoside reverse transcriptase inhibitors (NRTIs; abacavir with lamivudine or tenofovir disoproxil fumarate with emtricitabine) and an integrase strand transfer inhibitor, such as dolutegravir, elvitegravir, or raltegravir; a nonnucleoside reverse transcriptase inhibitor (efavirenz or rilpivirine) or a boosted protease inhibitor (darunavir or atazanavir) are rcommended for initial regimens (Günthard et al., 2014).

Fostemsavir (entry inhibitor via gp120) and PRO140 (CCR5 monoclonal antibody) are the two additional viral entry inhibitors with novel mechanisms of action that are currently in phase 2 trials (Gravatt et al., 2017). A phase 3 study is currently ongoing (NCT02362503) to determine if fostemsavir is an effective treatment for patients with multidrug-resistant HIV. PRO140 (CytoDyn) is a humanized CCR5 monoclonal antibody with antiviral activity against CCR5-tropic HIV. Based on new evidence assessing benefits and risks, the WHO recommends the use of the HIV drug dolutegravir (DTG) as the preferred first-line and second-line treatment for all populations, including pregnant women and those of childbearing potential (WHO 2019).

Antiretroviral drugs for HIV infection has been classified into the following categories: Multi-class Combination Products, Nucleoside Reverse Transcriptase Inhibitors (NRTIs), Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs),Protease Inhibitors (PIs), Fusion Inhibitors,Entry Inhibitors—CCR5 co-receptor antagonist and HIV integrase strand transfer inhibitors.

Herbal Medicine In The Treatment Of HIV/AIDS: The use of herbal medicine is increasingly becoming more popular in many countries (Sabde et al., 2011). This practice has continued to be a main source of health care in the rural communities especially in developing

countries, since modern medicine has not been able to reach the majority of the populace. In Africa, traditional herbal medicines are often used as primary treatment for HIV/ AIDS and for HIV-related problems including dermatological disorders, nausea, depression, insomnia and weakness. In North America, commonly used herbal dietary supplements have been found to impede on ARV drug effectiveness. Specifically, garlic supplements (Allium sativum) and St John's Wort (Hypericum perforatum) have been shown to have detrimental effects on the plasma concentrations of saquinavir and indinavir (Piscitelli et al., 2002).

Plants, produce numerous secondary metabolites as evolutionary responses to infections by fungi, nematodes, and other organisms, to avoid herbivory, and to complete for light and space, such as phenolics, glycosides, alkaloids, coumarins, terpenoids, essential oils and peptides. These metabolites have been identified with different biological activities. Some of them play an important role in immune system enhancement, exhibiting antiviral potential, including viral infections associated with Human Immunodeficiency Virus type 1 (HIV-1) and 2 (HIV-2) as genetic variabilities. An increasing number of patients with HIV infection cannot use the currently approved anti-HIV drugs including the reverse transcriptase and protease inhibitors, due to the adverse reactions, particularly liver diseases, that have been reported for antiretroviral drugs.

Some Chinese herbal preparation which consists of 14 plants (*Coptis chinensis, Jasminum officinale, Wolfiporia extensa, Sparganium stoloniferum, Polygonatum odoratum*, and *Scrophularia buergeriana* was investigated during 24 weeks and observed to have increased plasma CD4 count and also showed inhibition of HIV growth.

Table 1a. Plants with proven anti-HIV activity

S. no	Plant Name	Part of the Plant	Family	Assay	Reference
1	*Aegle marmelos*	Leaves Fruits	Rutaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
2	*Adhatoda vasica*	Leaves	Acanthaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
3	*Allium sativum*	Bulbs	Amaryllidaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
4	*Alstonia scholaris*	Stem bark Leaves	Apocynaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
5	*Argemone mexicana*	Leaves	Papaveraceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
6	*Asparagus racemosus*	Roots	Asparagaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
7	*Aconitum kusnezoffii*	Aerial	Ranunculaceae	MT-4 cell Assay	L M Bedoya, S Sanchez-Palomino, M J Abad et al., 2001
8	*Anemarrhena asphodeloides*	Rhizoma	Liliaceae	MT-4 cell Assay	Bahare Salehi, Nanjangud V. Anil Kumar, Bilge Sener et al., 2015
9	*Angelica sinensis*	Root	Umbelliferae	MT-4 cell Assay	Carolyn Williams-Orlando., 2017

Continue Table 1

10	*Artemisia caruifolia*	Whole plant	Asteraceae	MT-4 cell Assay	Chao-Mei MA, Norio Nakamura, Masao Hattori., 2001
11	*Andrographis Paniculata*	Leaves	Acanthaceae	MT-4 cell Assay	Mayur M Uttekar, Tiyasa Das, Rohan S Pawar et al., 2012
12	*Azadirachta indica*	Leaves	Meliaceae	Syncytium reduction assay, ELISA, Anti-HIV-1 RT inhibitory activity	David, Pedroza-Escobar Benjamín, Serrano-Gallardo Luis Delia et al., 2017
13	*Areca Catechu*	Seed	Piperaceae	-	Senthil Amudhan, V Hazeena Begum, K. B. Hebba, 2019
14	*Alchornea laxiflora*	Leaf, root, stem	Euphorbiaceae	HIV-1 Integrase inhibitory activity, Cytotoxicity activity	fD.Mnkandhla, M Issacs, F.M. Muganza et al., 2019
15	*Butea monosperma*	Roots Stem Bark	Leguminosae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
16	*Betula pubescens*	Bark	Betulaceae	anti-HIV-1 integrase assay	Prapaporn Chaniad , Teeratad Sudsai, Abdi Wira Septama et al., 2019
17	*Cassia occidentalis*	Leaves	Fabaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
18	*Catharanthus roseus*	Leaves	Apocynaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
19	*Cissampleos parriera*	Aerial part	Menispermaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
20	*Colchicum luteum*	Bulbs	Colchicaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
21	*Coleus forskohlii*	Aerial part	Lamiaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
22	*Cryptocarya chinensis*	Wood	Lauraceae	HIV growth inhibition assay	Tian-Shung Wu, Chung-Ren Su, Kuo-Hsiung Lee, 2012
23	*Coccinium fenestratum*	Stem bark	Menispermaceae	Integrase and Protease Inhibitor assay	J.J. Magadulai, H.O. Suleiman., 2010
24	*Calophyllum inophyllum*	Bark	Guttiferae	RT Inhibition assay	J.J. Magadulai, H.O. Suleiman., 2010
25	*Cinnamomun aromiticum*	Bark	Lauraceae	MT-4 cell Assay	Franklin Nyenty Tabe, Nicolas Yanou Njintang, Armel Hervé Nwabo Kamdje et al., 2015
26	*Cynanchum chinense*	Root	Asclepiadaceae	MT-4 cell Assay	Jian Tao, Jing Yang, Chaoyin Chen et al., 2011
27	*Cynomorium songaricum*	Stem	Cynomoriaceae	MT-4 cell Assay	Suvdmaa Tuvaanjav, Han Shuqin, Masashi Komata et al., 2016

Continue Table 1

28	*Dracocephalum rupestre*	Whole plant	Labiatae	MT-4 cell Assay	Qi Zeng, Hui-Zi Jin, Jiang-Jiang Qin et al., 2010
29	*Dryopteris crassirhizoma*	-	Aspidiaceae	MT-4 cell Assay	Ji Suk Lee, Hirotsugu Miyashiro, Norio Nakamura et al., 2008
30	*Embelica ribes*	Fruits	Primulaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
31	*Embellica officinalis*	Fruits	Phyllanthaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
32	*Erodium stephanianum*	Whole plant	Geraniaceae	MT-4 cell Assay	Chao-mei Ma, Norio Nakamura, Hirotsugu Miyashiro, 2002
33	*Eugenia jambolona*	Bark	Myrtaceae	-	Richa Sood, D Swarup, S Bhatia, D D Kulkarni et al., 2012
34	*Garcinia indica*	Leaves	Clusiaceae	MT-4 cell Assay	J.J. Magadulai, H.O. Suleiman., 2010
35	*Garcinia cambogia*	Leaves	Clusiaceae	Integrase and Protease Inhibitor assay	J.J. Magadulai, H.O. Suleiman., 2010

Table 1b. Plants with proven anti-HIV activity

S. no	Plant Name	Part of the Plant	Family	Assay	Reference
1	*Gentiana scabra*	Root	Centianaceae	MT-4 cell Assay	Bahare Salehi, Nv Anil, Bilge Sener et al., 2018
2	*Gossampinus malabarica*	Flower	Bombacaeae	MT-4 cell Assay	J A Wu, A S Attele, L Zhang et al., 2001
3	*Gymnadenia conopsea*	Root	Orchidaceae	MT-4 cell Assay	Xiaofei Shang, Xiao Guo,Yu Liu et al., 2017
4	*Glycyrrhiza glabra*	Glcyrrhyrine	Fabaceae	OKM-1, MT-4 cells	Cristina Fiore, Michael Eisenhut, Rea Krausse et al., 2008
5	*Glycyrrhiza glabra*	Roots	Fabaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
6	*Gentiana scabra*	Root	Centianaceae	MT-4 cell Assay	Bahare Salehi, Nv Anil, Bilge Sener et al., 2018
7	*Lygodium japonicum*	Spore	Schizaeaceae	MT-4 cell Assay	Xavier-ravi Baskaran, Antony-varuvel Geo Vigila, Shou-zhou Zhang et al., 2018
8	*Madhuca indica*	Bark	Sapotaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
9	*Morinda citrifolia*	Leaves	Rubiaceae	MT-4 cell Assay	P. Selvam, N. Murugesh, M. Witvrouw et al., 2009
10	*Moringa oleifera*	Leaves	Moringaceae	Vector based antiviral assay	Nworu CS, Ezeifeka GO, Ebele Okoye et al., 2013

Continue Table 1b

11	*Myrianthus holstii*	Root	Urticaceae	Synctia Formation assay	Michael J. Currens, Lewis K. Pannell, and Michael R. Boyd et al., 2000
12	*Ocimum sanctum*	Leaves	Lamiaceae	RT Inhibition assay, Gp120 Binding Inhibition assay	Kun Silprasit, Supaporn Seetaha, Parinya Pongsanarakul et al., 2011
13	*Oldenlandia diffusa*	Whole plant	Rubiaceae	MT-4 cell Assay	Bahare Salehi, Nanjangud V. Kumar, Anil Bilge Sener et al., 2018
14	*Polygonum divaricatum*	Whole plant	Polygonaceae	MT-4 cell Assay	Yu Zhong, Yoshiyuki Yoshinaka, Tadahiro Takeda et al., 2005
15	*Papaver somniferum*	Seeds	Papaveraceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
16	*Piper longum*	Fruit	Piperaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
17	*Phyllanthus amarus Schum*	Leaves	Phyllanthaceae	RT assay	F Notka, G R Meier, Ralf Wagner, 2003
18	*Phyllanthus emblica*	Fruit	Phyllanthaceae	p24 production assay	M Estari, L Venkanna, D Sripriya et al., 2012
19	*Pelargonium sidoides*	Root	Geraniaceae	HIV-1-cell attachment assays	Markus Helfer, Herwig Koppensteiner, Martha Schneider et al., 2014
20	*Rubia cordifolia*	Roots	Rubiaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
21	*Rhaponticum uniflorum*	Root	Compositae	MT-4 cell Assay	Hai-Li Liu,Yue -Wei Guo, 2008
22	*Rubia cordifolia* L	Root	Rubiaceae	MT-4 cell Assay	Yuanyuan Sun, Xuepeng Gong, Jia Y Tan et al., 2016
23	*Rauwolfia serpentina*	Roots	Apocynaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
24	*Papaver somniferum*	Seeds	Papaveraceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
25	*Piper longum*	Fruit	Piperaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
26	*Phyllanthus amarus Schum*	Leaves	Phyllanthaceae	RT assay	F Notka, G R Meier, Ralf Wagner, 2003
27	*Salacia oblonga*	Leaves	Celastraceae	Integrase and Protease Inhibitor assay	J.J. Magadulai, H.O. Suleiman., 2010
28	*Salvia miltiorrhiza*	Roots	Lamiaceae	MTT assay, Virus neutralization assay	Ibrahim S Abd-Elazem, Hong S Chen, Robert B Bates et al., 2002
29	*Silybum marianum*	-	Asteraceae	MT-4 cell Assay	Ching-Hsuan Liu, Alagie Jassey,
30	*Scorzonera glabra*	Root	Compositae	MT-4 cell Assay	Hsin-Ya Hsu et al., 2019 Chao-mei Ma, Norio Nakamura, Hirotsugu Miyashiro et al., 2002

Continue Table 1b

31	*Scutellaria barbata*	Whole plant	Portulacaceae	MT-4 cell Assay	Zi-Long Wang, Shuang Wang,
32	*Stellera chamaejasme*	Root	Thymelaeaceae	MT-4 cell Assay	Yi Kuang et al., 2018 Min Yan, Yan Lu, Chin-Ho Chen et al., 2015
33	*Stephania cepharantha*	Root, Tuber	Menispermaceae	MT-4 cell Assay	Chao-mei Ma 1, Norio Nakamura, Hirotsugu Miyashiro et al., 2002
34	*Sterculia scaphigera*	Seed	Sterculiaceae	MT-4 cell Assay	Moshera Mohamed El-Sherei, Alia Ragheb, Mona Kassem et al.,2016
35	*Tinospora cordifolia*	Stem bark	Menispermaceae	p24 antigen assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
36	*Terminalia sericea*	Leaves	Combretaceae	MTT assay	M A Chauke, L J Shai, M A Mogale et al., 2016
37	*Withania somnifera*	Roots	Solanaceae	p24 antigen assay, Gp120 Binding Inhibition assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011
38	*Withania somnifera*	Roots	Solanaceae	p24 antigen Gp120 assay, Binding Inhibition assay	Sudeep Sabde, Hardik S. Bodiwala, Aniket Karmase et al., 2011

Cytotoxicity of Anti-HIV Phytochemicals: Cytotoxic evaluation is very important and integral part of research involving discoveries of new and potent antiviral drugs. A novel formulation with potent antiviral activity have to be proven as not having any toxicity effects and cytotoxicity assays in a suitable cell culture system are only a part of primary step in this direction. For the purpose of testing, different plants active principals have to be extracted with suitable solvents. The list of commonly used solvents for extraction purpose is summarized in Table 2. Treating cells with these phytochemicals can result in a variety of cell fates. The cells may undergo necrosis, in which they lose membrane integrity.

Table 2. Solvents used for active components extraction

Water	Ethanol	Methanol	Chloroform	Di-chloro methanol	Ether	Acetone
Anthocyanins	Tannins	Anthocyanins	Terpenoids	Terpenoids	Alkaloids	Flavanols
Starches	Polyphenols	Terpenoids	Flavonoids		Terpenoids	
Tannins	Polyacetylenes	Saponins			Coumarins	
Saponins	Flavanol	Tannins			Fatty acids	
Terpenoids	Terpenoids	Xanthophyllines				
Polypeptides	Sterols	Totarol				
		Lactones				

Cytotoxicity can also be monitored using the MTT or MTS assay. This assay measures the reducing potential of the cell using a colorimetric reaction. Viable cells will reduce the MTS reagent to a colored formazan product. Tetrazolium salts are reduced only by metabolically active cells. Thus, 3-(4, 5-dimethylthiazol-2- yl)-2, 5-diphenyltetrazolium bromide (MTT) can be reduced to a blue colored formazan32. A similar redox-based assay has also been developed using the fluorescent dye, resazurin. In addition to using dyes to indicate the redox potential of cells in order to monitor their viability, researchers have developed assays that use ATP content as a marker of viability (Riss et al., 2004). Adenosine triphosphate (ATP) that is present in all metabolically active cells can be determined in a bioluminescent measurement. The bioluminescent method utilizes an enzyme, luciferase, which catalyses the formation of light from ATP and luciferin. The emitted light intensity is linearly related to the ATP concentration (Weyermann et al.,2005). Neutral red (3- amino-m-dimethylamino-2-methylphenazine hydrochloride) has been used previously for the identification of vital cells in cultures. This assay quantifies the number of viable, uninjured cells after their exposure to toxicants; it is based on the uptake and subsequent lysosomal accumulation of the supravital dye, neutral red. Quantification of

the dye extracted from the cells has been shown to be linear with cell numbers, both by direct cell counts and by protein determinations of cell populations (Weyermann et al., 2005).

Future Perspectives: In vitro studies of many plant phytoconstituents can be evaluated for various anti-viral activities including anti-HIV activity and COVID-19. Further studies can be carried out to know the mechanism of drug inhibition in virus. Synthetic drugs are proved to cause side effects. However, more exploratory research to prove the efficacy of medicinal plants including plant - drug interactions and their mechanism of action has to be explored so that plant compounds can be used to treat various viral infections including deadly COVID-19.

CONCLUSION

Many plant species have been investigated for anti-HIV potential and has shown promising activity. Azidothymidine, the first drug that was approved in the fight against AIDS in the 1980s, still a main component in the medication mix commonly prescribed to HIV patients today. But new research have found a plant-derived chemical compound that is much more effective than azidothymidine. The chemical compound is called "patentiflorin A" and is derived from a medicinal plant found in East Asia: Justicia gendarussa. Hence, plant based source drugs are non-toxic and work effectively unlike synthetic drugs. Many synthetic medicines are being used in the treatment of AIDS. Various medicinal plants or plant-derived natural products has offered alternatives to expensive medicines in future.

REFERENCES

Akkouh O, Ng TB, Singh SS, Yin C, Dan X, Chan YS et al., (2015) Lectins with anti-HIV activity: a review Molecules 20, 648668

Akram Muhammad, Tahir Mahmood, Munir et al., (2017) Antiviral potential of medicinal plants against HIV, HSV, influenza, hepatitis, and coxsackievirus: A systematic review. Phytotherapy Research 1–12.

Awah FM, Uzoegwu PN, Ifeonu P (2011) In vitro anti-HIV and immunomodulatory potentials of *Azadirachta indica* (Meliaceae) leaf extract African Journal of Pharmacy and Pharmacology Volume 5(11) pp. 1353-1359

Blanco JL, Whitlock G, Milinkovic A, Moyle G (2015) HIV integrase inhibitors: A new era in the treatment of HIV Expert Opinion Pharmacotherapy 16, 313–1324.

Behbahani M, (2014) Evaluation of anti-HIV-1 activity of a new iridoid glycoside isolated from *Avicenna marina*, in vitro International Immunopharmacology 23, 262–266.

Chang YC, Hsieh PW, Chang FR, Wu RR, Liaw CC, Lee KH, Wu YC (2003) Two new protopines argemexicaines A and B and the anti-HIV alkaloid 6-acetonyldihydrochelerythrine from formosan *Argemone mexicana*. Planta Medica 69:148–152

Chukwujekwu JC, Ndhlala AR, de Kock CA, Smith PJ. Van Staden J, (2014) Antiplasmodial, HIV-1 reverse transcriptase inhibitory and cytotoxicity properties of *Centratherum punctatum* Cass. and its fractions South African Journal of Botany 90, 17–19.

Esposito F, Mandrone M, Del Vecchio C, Carli I, Distinto S, Corona A, Lianza M, Piano D, Tacchini M, Maccioni E et al.,(2017) Multi-target activity of *Hemidesmus indicus* decoction against innovative HIV-1 drug targets and characterization of lupeol mode of action. Pathogens and Disease 75.

Gambari Raberto, Lampronti Ilaria (2006) Inhibition of immunodeficiency type-1 virus (HIV-1) life cycle by medicinal plant extracts and plant-derived compounds Elsevier Volume 2 Pages 299-311.

Gravatt L A H, Leibrand CR, Patel S, McRae M (2017) New drugs in the pipeline for the treatment of HIV: A review. Current Infectious Disease Reports Sep 19;19(11):42

Günthard HF, Aberg JA, Eron JJ, Hoy JF, Telenti A, Benson CA, Burger DM, Cahn P, Gallant JE, Glesby MJ (2014) Antiretroviral treatment of adult HIV infection: recommendations of the International Antiviral Society—USA panel. Journal of the American Medical Association 312, 410–425

Helfer M, Koppensteiner H, Schneider M, Rebensburg S, Forcisi S, Muller C et al., (2014) The root extract of the medicinal plant *Pelargonium sidoides* is a potent HIV-1 attachment inhibitor PLoS One 9.

Jadhav AN, Bhutani KK (2006) Steroidal saponins from the roots of *Asparagus adscendens* Roxb and *Asparagus racemosus* Willd. Indian J Chem Indian Journal of Chemistry 45B:1515–1524

Kapewangolo P, Knott M, Shithigona REK, Uusiku SL, Kandawa-Schulz M, (2016) In vitro anti-HIV and antioxidant activity of *Hoodia gordonii* (Apocynaceae), a commercial plant product BMC Complementary and Alternative Medicine 16, 411.

Kurapati KRV, Atluri VS, Samikkannu T, Garcia G, Nair MP (2014) Natural products as anti-HIV agents and role in HIV-associated neurocognitive disorders (hand): A brief overview. Frontiers in Microbiology 6:1444

Laila Umme, Akram Muhammad, Shariati Ali Mohammad et al.,(2019) Role of medicinal plants in HIV/AIDS therapy Clinical and Experimental Pharmacology and Physiology Volume 46 Issue12 Pages 1063-1073

Ma C Nakamura, N Miyashiro, H Hattori, M Komatsu, K Kawahata et al., (2002) Screening of Chinese and Mongolian herbal drugs for anti-human immunodeficiency virus type 1(HIV-1) activity Phytotherapy Research 16, 186189

Maartens G, Celum C, Lewin SR (2014) HIV infection: Epidemiology, pathogenesis, treatment, and prevention Lancet 384, 258–271

Modi M, Goel T, Das T, Malik S, Suri S et al., (2013) Ellagic acid & gallic acid from *Lagerstroemia speciosa* L. inhibit HIV-1 infection through inhibition of HIV-1 protease & reverse transcriptase activity Indian Journal

of Medical Research 137, 540548
Mugomeri Eltony, Chatanga Peter, Chakane Ntema (2016) Medicinal Herbs used by HIV-Positive people in Lesotho African Journal of Traditional, Complementary and Alternative Medicines 13(4): 123–131.
Mukhtar Muhammad, Arshad Mohammad, Ahmad Mahmood et al.,(2013) Antiviral Potential of Medicinal Plants : An overview International Research Journal of Pharmacy 4(6).
Narayan Chaitra, Vittal Rai Ravishankar, Tewtrakul Supinya (2011) A Screening strategy for Selection of Anti-Hiv-1 Integrase and Anti-Hiv-1 Protease Inhibitors from Extracts of Indian Medicinal Plants International Journal of Phytomedicine Volume 11 No 2.
Noundou Siwe, Ndintehde DT, Olivier DK, Mnkandhla D, Isaacs M et al., (2018) Biological activity of plant extracts and isolated compounds from *Alchornea laxiflora*: Anti-HIV, antibacterial and cytotoxicity evaluation South African Journal of Botany-02111 No of Pages 6
Nworu CS, Okoye EL, Ezeifeka GO, Esimone CO (2013) Extracts of Moringa Oleifera Lam. showing Inhibitory Activity against early steps in the Infectivity of HIV-1 Lentiviral particles in a viral vector-based screening African Journal of Biotechnology Volume. 12(30) pp. 4866-4873
Olatunya OS, Olatunya AM, Anyabolu HC, Adejuyigbe EA, Oyelami OA (2012) Preliminary Trial of Aloe Vera Gruel on HIV Infection Journal of Alternative and Complementary Medicine Sep;18(9):850-3.
Palshetkar Aparna, Pathare Navin, Jadhav Nutan et al.,(2020) In vitro anti-HIV activity of some Indian medicinal plant extracts BMC Complementary Medicine and Therapies 20, Article number: 69
Patridge E, Gareiss P, Kinch MS, Hoyer D (2016) An analysis of FDA-approved drugs: natural products and their derivatives Drug Discovery Today 21, 204207
Piscitelli SC, Burstein AH, Welden N, Gallicano KD, Falloon J (2002) The Effect of Garlic Supplements on the Pharmacokinetics of Saquinavir Clinical Infectious Diseases Jan 15;34(2):234-8.
Rege A Anuya, Ambaye Y Ramakrishna, Deshmukh A Ranjana (2010) In-Vitro Testing of Anti-HIV Activity of Some Medicinal Plants Indian Journal of Natural Products and Resources. Volume 1(2) June pp. 193-199
Rege AA, Chowdhary AS, (2013) Evaluation of mangrove plants as putative HIV-protease inhibitors Indian Drugs 50, 41–44.
Riss TL, Moravec RA (2004) Use of multiple assay endpoints to investigate the effects of incubation time, dose of toxin, and plating density in cell-based cytotoxicity assays. ASSAY and Drug Development Technologies Feb;2(1):51-62.
Sabde Sudeep, Bodiwala Hardik, Karmase Aniket et al Anti-HIV activity of Indian medicinal plants (2011) Journal of Natural Medicines, 65:662–669
Salehi Bahare, Nanjangud V. Anil Kumar, Cener Bilge, et al.,(2018) Medicinal Plants Used in the Treatment of Human Immunodeficiency Virus International Journal of Molecular Sciences May;19(5): 1459.
Shan Y, Wang X, Zhou X, Kong L, Niwa M (2007) Two minor diterpene glycosides and an eudesman sesquiterpene from *Coleus forskohlii*. Chemical and Pharmaceutical Bulletin 55:376–381
Singh IP, Bodiwala HS (2010) Recent advances in anti-HIV natural products. Natural Product Reports 27, 1781
Tabe Franklin Nyenty, Yanou Nicolas Njintang, Kamdje Nwabo et al., (2015) Oxidative Role of HIV/AIDS: Antiretroviral Drugs and Medicinal Plants with Anti-HIV Activity Journal of Diseases and Medicinal Plants 1(5): 68-75
Thenin-Houssier S, Valente ST (2016) HIV-1 capsid inhibitors as antiretroviral agents. Current HIV Research 14, 27028
Trivedi Jay, Tripathi Anjali, Chattopadhyay Debprasad, Mitra Debashis (2019) Chapter 11 - Plant-Derived Molecules in Managing HIV Infection Advancements in Herbal Products as Novel Drug Leads pp 273-2983
Visintini María, Redko Flavia, Muschietti Liliana et al., (2013) In vitro antiviral activity of plant extracts from Asteraceae medicinal plants Virology Journal Volume 10.
Wang B, Ge L, Huang W, Zhang H, Qian H, Li J et al.,(2010) Synthesis and preliminary anti-HIV activities of andrographolide derivatives. Journal of Medicinal Chemistry (Los Angeles) 6, 252258
Weyermann J, Lochmann D, Zimmer A (2005) A practical note on the use of cytotoxicity assays International Journal of Pharmaceutics Jan 20;288(2):369-76.
World Health Organization (WHO) [(accessed on 25 December 2019)];2019
Wu TS, Su CR, Lee KH (2012) Cytotoxic and anti-HIV phenanthroindolizidine alkaloids from *Cryptocarya chinensis* Natural Product Communications 7, 725727
Zhou P, Takaishi Y, Duan H, Chen B, Honda G, Itoh M, Takeda Y, Kodzhimatov OK, Lee K-H (2000) Coumarins and bicoumarin from *Ferula sumbul*: anti-HIV activity and inhibition of cytokine release Phytochemistry 53:689–697
Zou Wen, Liu Ying, Wang Jian, Li Hongjuan, Liao Xing (2012) Traditional Chinese Herbal Medicines for Treating HIV Infections and AIDS. Evidence Based Complement Alternate Medicine 2012:950757 8 pages.

Ensemble and Feature Selection-based Intrusion Detection System for Multi-attack Environment

S. R. Khonde
Department of CSE,
Sathyabama Institute of Science and Technology,
Chennai, India
khonde.shraddha@gmail.com

Dr. V. Ulagamuthalvi
Department of CSE,
Sathyabama Institute of Science and Technology,
Chennai, India
ulagamv@gmail.com

Abstract – **In the world of internet most challenging issue is to maintain confidentiality and integrity of data. Network provides poor performance in terms of security when it gets compromised by some intruder. Intruders insert intrusions in the network to compromise system and get access to data. To compromise network intrusions are inserted in the form of attack. An attack enters into a network as a normal traffic and then tries to compromise devices connected into network. Now day's intruders are using multi-attacks to break security of network. To detect such type of activities intrusion detection system (IDS) is used. IDS keeps eye on each packet entering into the network. If any abnormal behavior is identified an alert is generated for administrator. However IDS are able to detect only single attack at a time happening on the network. To handle modern era multi-attack improvement in IDS is required. In this paper an improved IDS is presented which can handle multi-attack with the help of machine learning classifiers. Ensemble approach using majority voting algorithm is used for improving performance. To reduce training and computation time feature selection technique is used. Performance evaluation is done using precision, accuracy and false alarm rate. Dataset used for experiments is NSL-KDD. Comparison of individual classifier is done with ensemble approach. Results show that ensemble approach shows reduction of 0.05% in false alarm rate and 2% improvement in accuracy of IDS. Precision of ensemble is also increased by 0.01. Ensemble approach and feature selection provides improved performance of IDS.**

Keywords: multi-attack, network security, intrusion detection system, random forest, decision tree, naïve bayes, feature selection

I. INTRODUCTION

Rapid increase in usage of internet takes users towards the new era and architecture of network. This modern network can also be called as a multimedia era which allows users to transfer multimedia data consist of text, audio, video etc. This flexibility to transfer any type of data in any corner of world using internet raise question of security. Security of data is utmost important when it comes to network or internet. To provide security to data passing through network and secures all systems connected in network we require a device which will provide security. A device which helps us in providing security to network is called intrusion detection system. This system can also use as a software application to provide security to data. IDS monitor all the activities of network. Each data entering and leaving network is checked by IDS. Data enters in the network in form of packet. Packets are analysed by IDS to check whether it is a normal packet or malicious packets. Intruders make use of attacks to compromise network. An attack can be defined as a threat to system which hides from IDS and remain present in network. This can be a software application or program which IDS not able to detect. This attack tries to compromise system and network to do various malicious activities and harm data. Data confidentiality and integrity is not maintained by network once it got attack activated inside it. Two types of attacks can happen on network as active and passive attacks. Active attacks captures data from network modifies it and then passed it further. This attack can modify, delete as well as edit data that is capture. This type of attack is very dangerous if enters in the network. Passive attack only remains in network or system to keep eye on activities happening in network to receive confidential information. Another emerging attack is multi-attack which is created using various combinations of attacks and hard to detect. Detection is done in two approaches as given below

1.1 Signature based IDS

Signature or misuse based IDS make use of signatures to detect patterns of attacks. All the previously happened attacks foot prints are saved in the dataset. These are the standard datasets consist of signatures of various attacks. Whenever any packet enters into the network signature based IDS matches its pattern with the existing signatures it match found then that attack is considered as malicious packet and alert is generated for administrator. If match does not found then the packet is allowed to pass in the network to reach client. Limitation of these IDS is it will not identify novel attacks. Novel attacks are type of attacks whose signature is not available in the dataset.

1.2 Anomaly based IDS

Novel attacks which are not detected by signature based IDS is identified by this type of IDS. Anomalies are detected from the packets entering in the system. This new IDS does not use any type of signatures. It only freezes the normal behaviour of the traffic. Normal behaviour activities are declared by classifier. Then IDS will analyse each packet based on this behaviour. If any change in behaviour of traffic is identified it is considered as a malicious packet. Anomaly based IDS provide detection of various modern era novel attacks which signature based IDS cannot do. Limitation of these IDS is it generates more false predictions as compare of signature based IDS. Reason behind it is slight change in behaviour is also consider as attack in this IDS. This is one of the popularly used IDS now days.

978-1-7281-9180-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

Both types of IDS do attack detection of a single attack at a time. This paper presents improved IDS which will be able to identify or detect multi-attack at a time. Multi-attack is defined as multiple attacks are happening on the network at the same time or with the difference of fraction of seconds. Multi-attack can also be any new attack created by combining two or more existing attacks. All these types of multi-attacks with the regular single attacks can be identified by IDS. Machine learning classifiers are used for classifying attack and normal traffic. Dataset used for experiments is standard NSL-KDD intrusion detection dataset. This dataset consist of 38 different types of attack signatures. To detect multi-attack, signatures for that is created and inserted into same dataset. To train various classifiers rather than using all features of KDD that is 41, feature reduction is done to reduced number of features. This new IDS provides improved accuracy and detection of attack with reduced False alarm rate (FAR).

II. Literature Survey

Network security can be improved by using machine learning algorithms in IDS. These algorithms provide better accuracy when used with standard dataset. Various datasets available are KDD99, NSL-KDD, DARPA, UNSW-NB15, CIC-IDS2017 and many more. These all datasets are used by researchers to implement IDS for attack detection. Machine learning algorithms are divided into three categories as unsupervised, semi-supervised and supervised algorithms. IDS with ensemble classifiers random forest, naïve bayes and average one-dependence estimator can be used [1]. Experiments performed on Kyoto dataset with 90.51% accuracy and 0.14 FAR. Another ensemble approach is presented where core vector machine are used to identify all categories of attacks of KDD99 dataset [2]. For each category one CVM is implemented and tested for attack detection. Chi square test is used for feature selection. Results prove that this approach provides good efficiency and less computational time. Identification of binary and multi-class classifications approach is used in [3]. Features are selected using rankers based attribute evaluation technique. Experiments are done on KDD dataset and give binary class accuracy as 99.72% and for multi-class it is 99.68%. Feature selection plays an important role in the performance of IDS [4]. Various feature selection techniques as chi square, gain ratio, cfs and information gain on ensemble and single classifiers can be used. NSL-KDD dataset is used for experiments. Ensemble adaboost methods show better accuracy as compared to bagging and stacking. Framework for feature selection is presented by [5]. It uses two feature selection techniques as correlation and hypothetical testing-based feature selection. Both techniques are tested on KDD99 and NSL-KDD dataset. High accuracy is shown by NSL-KDD. Ensemble method predictions are more accurate as compared to single classifier [6]. Authors also elaborate various ensemble techniques which can be used which implementation of IDS. Feature selection based on genetic algorithms is used for IDS [7]. Features reduced up to 15 of NSL-KDD dataset with decision tree. Results show 99.71% accuracy which is higher than other classifiers. Wrapper feature selection technique is most used one now days [8]. Extreme machine learning is used to test NSL-KDD dataset with reduces features. Accuracy of 80.15% is obtained using this wrapper selection technique. In [9] comparison of various machine learning algorithms according to their performance is elaborated. Classifiers are tested on KDD-CUP dataset to provide comparison between them. An ensemble approach of genetic algorithm and support vector machine to improve performance of IDS is given in [10]. Fitness function is used on genetic algorithm to reduce number of features and tested on SVM to reduce its error rate and increases positive rate. Various approaches to implement IDS like ensemble and hybrid classifiers are used by researchers with homogeneous and heterogeneous classifiers [11]. Some privacy issues can be faced while IDS implementation [12]. Performance of various classifiers is compared with help of precision and accuracy parameters. To reduce FAR in IDS various techniques are presented in [13]. IDS for wired mesh networks using genetic algorithm and feature selection is explained by [14]. Multiple SVM classifiers are used to improve accuracy. Flow based IDS using packet evaluation can be used for improving IDS performance [15]. A novel framework is proposed which use SVM ensemble with feature augmentation [16]. Result proves better performance in terms of accuracy, FAR, training time and detection rate. Hybrid approach for ensemble classifiers is elaborated in [17]. In [18] a novel network intrusion detection system based on KELM ensemble approach is explained. This approach provides better accuracy than traditional NIDS. Performance test of Bayesian network and CART algorithms with feature selection techniques is done [19]. Ensemble of these two classifiers gives more performance than individual classifier. Feature selection technique to improve performance of IDS in binary and multiclass classifiers with hybrid model is explained in [20]. It provides accuracy of 99.77% for binary and 99.63% for multiclass detection. Feature selection with random forest gives better performance as presented in [21]. Random mutation hill climbing based wrapper feature selection also helps in improving IDS performance [22]. Comparison of filter and wrapper feature selection proves that both methods are efficient and helps IDS to perform better [23]. Performance of IDS is depend of classifier, feature selection and datasets used. Survey for all standard dataset variables for network IDS is provided in [24]. Update in dataset require sometime to improve IDS performance. As per observations in literature survey presented above we can say that IDS provides improved performance if classifiers are used in ensemble approach rather than individually. Feature selection help to reduce computation and training time of classifier. In proposed IDS multiple classifiers with ensemble approach is used. It makes use of feature selection techniques. As none of the benchmarked dataset includes signature of multi-attack it is generated and updated in NSL-KDD dataset.

III. METHODOLOGY

Most of IDS available today is used for single attack detection. Proposed IDS can detect single as well as multi-attacks. IDS make use of machine learning algorithms like decision tree (DT), naïve bayes (NB) and random forest (RF). These three classifiers are used in ensemble bagging for implementation of IDS. Dataset used is NSL-KDD, which will be modified as per requirement of multi-attack detection. Signatures of multi-attacks will be generated and stored in dataset. To reduce training and computation time various feature selection techniques are used in pipeline so that features can be reduced and detection rate can be increased. Feature reduction techniques used are gain ratio, information gain and correlation coefficient method.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

3.1. Customization of dataset

The standard dataset used to implement proposed system is NSL-KDD. This dataset is a standard dataset used for intrusion detection. NSL-KDD is the better version of KDD99 dataset. This dataset does not consist of any redundancies or duplications of records. Total number of attributes in this dataset is 42. This dataset has training and testing parts which are used to train classifiers and test it for attack detection. More than lakh instances are available in the dataset where total 38 types of attacks signatures are available. In total 41 attributes are used for attack detection and last attribute is used to label it. Label assigned is "normal" and "attack". Various categories of attacks are covered in this dataset as Denial of service, Root to local (r2l), probe and User to root (u2r). In total 38 attacks signatures are available in these four categories. This dataset does not contain signature of multi-attack.

To create a signature of a multi-attack, scripts for multi-attack is written and penetrated into the network. Scripts consist of combination of multiple attacks which will create signature of new attacks whose behaviour will be combination of behaviour of attacks used in scripts. This is an important step in customization of dataset. Fig. 1 shows process for dataset customization. New dataset will consist of signature of all 38 attacks plus signature of multi-attacks.

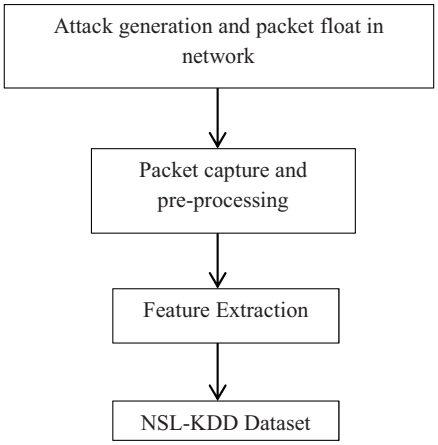


Fig. 1 Dataset Customization Process

Packets of multi-attacks are randomly generated and penetrated into the network along with the normal traffic. These packets were captured using packet analyser tool Ethereal. This tool allows us to capture and analyse all the packets from the traffic. This packets information is saved into a pcap file. This file is passed for pre-processing to remove all the noise and invalid data. Once the packets data is clean features are extracted form it according to feature selection techniques used. These extracted features are saved into the dataset. Customized dataset will consist of only features decided after feature selection. This dataset will consist of single as well as multi-attack signatures. Feature selected using feature selection technique is explained in next section.

3.2 Feature selection

Feature selection is used for reducing number of features from the dataset. It will help to reduce computational cost as well as training time for classifiers. Features should be reduced in manner that it will not affect performance or accuracy of the IDS. Detection rate provided by IDS with all features should be obtained after reducing features. Motto behind reducing it is only to decreases training and computation time. Various feature selection techniques are available for dimensionality reduction. Techniques used in proposed system are explained in further sections.

3.2.1 Correlation coefficient

This method used to show relation between variables. All features are related with each or not are obtained from this method. Relation between two variables can be positive or negative. Positive relation shows the strong relation between that two features and negative relation shows features are independent from each other. This relation measure is used to select strongly related features so that attack detection can be done with help of less number of attributes or features rather than using all features. In every dataset some features are totally independent what does not take part in attack detection process. Such attributes or features need to be removed while training. This method gives us updated attribute list which can be used for training classifier.

3.2.2 Entropy and Information Gain

In algorithms for tree like structure we need to calculate quality of each split done for finding the accuracy of the classifier. Information gain is used to check the accuracy of split used in classifiers to make classifications. This also shows the amount of information obtained from that attribute. This information is used to reduce features. The features having less information gain is removed from the feature list used to train classifiers.

3.2.3 Gain Ratio

It is the ratio of information gain occurs from the decision tree split information. To avoid limitation of information gain mostly gain ratio is used. It use to handle various features having multi-valued attributes. Though features are multi-valued gain ratio did not provide biased prediction. It is used to provide amount of gain obtained from features after considering its entropy or information gain.

All these three methods are tested on decision tree and random forest classifier using WEKA tool. This tool allows checking the performance of each classifier using a set of attributes. Classifier accuracy is obtained first with the help of all attributes that is 42 of NSL-KDD dataset. Then after using all method of feature selection accuracy is tested for each category of attack from dataset. Number of features reduced according to its gain value. The features having less value are removed to find new set of features. This reduces feature set is then used to train classifier. Same WEKA tool is used to check performance of classifiers using reduced dataset. Number of

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

features used to train using Gain ratio feature selection technique is 29 features, for information gain it is 28 features and for correlation coefficient 25 features are used. These all selection methods are tested individually for checking performance of classifier. To avoid limitation of each method and maintain accuracy pipelining is used for feature selection. Features set obtained from Gain ratio is passed as input to information gain and then to correlation coefficient. Finally we obtained only 25 features as a reduced set with maintained classifier accuracy. Pipelining of feature selection methods is shown in figure 2. Reduce features improves accuracy of all classifiers and performance can be maintained. Numbers shown in bracket is the number of features obtained from method after reduction.

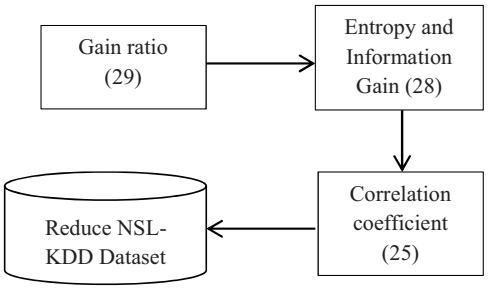


Fig. 2 Pipelining of feature selection techniques

List of features after feature selection is given below:

['src_bytes', 'dst_bytes', 'land', 'wrong_fragment', 'urgent','srv_count', 'serror_rate', 'srv_serror_rate', 'same_srv_rate','diff_srv_rate', 'srv_diff_host_rate', 'dst_host_diff_srv_rate','dst_host_same_src_port_rate', dst_host_srv_diff_host_rate','dst_host_serror_rate', dst_host_srv_serror_rate','dst_host_rerror_rate', dst_host_srv_rerror_rate', 'protocol_type','service', 'flag', 'xAttack']

These final 25 features are used to trained classifiers in proposed architecture. Feature selection is done by cross fold 10 validation using all above mentioned feature selection techniques. This reduced dataset will consist of more than a lakh records consisting of single and multi-attack signatures.

3.3. Classifiers

Machine learning algorithms are used as classifiers to implement IDS. Various machine learning algorithms are used to build a misuse based intrusion detection system. Various categories of algorithms machine learning have as supervised which works on labelled data, unsupervised works with unlabelled data and combination of both semi-supervised which works with labelled as well as unlabelled data. As proposed IDS is based on signatures, dataset consist of labelled data. Supervised machine learning classifiers are used to identify malicious activities. Classifiers used are explained below

3.3.1 Naïve Bayes

This classifier works on probability of attack occurrence. If some features are identified form malicious activities them this algorithms helps us to find probability for occurrence of other feature. Using this classifier we can identify probability of attack detection in the traffic. This classifier works on labelled data. This algorithm based on the principal of bayes theorem which is used to check the independence of all features. Training time of this classifier is very fast and it provides accurate prediction for classification of data. Naïve bayes classifier focuses on the independence of features of dataset used for attack detection.

3.3.2 Decision tree

Decision tree is a supervised learning classifier which works on labelled data. This classifier has a tree like structure which consists of root node, internal nodes, leaf nodes and edges. In the tree of this classifier root node is for the prediction of a class. Any test is perform on root node will have possibilities. All possibilities are known as internal nodes. For further classification and reach up to class test is done on internal node to get the choices. This will continue till you reach up to leaf node. Leaf node is used to assign labels to the data. The path from root node to leaf node is a single instant or record for the dataset. Final label is obtained from the leaf node. Edges connecting nodes are output of test done on nodes. Decision tree provides better accuracy for labelled dataset. Limitation of this dataset is it provides biased classification for multi—valued attributes.

3.3.3 Random forest

To remove biased prediction of decision tree we can use random forest classifier. It consists of various trees and each tree works as a decision tree. Accuracy of this classifier is based on number of trees used for identification. Random forest itself is an ensemble classifier with number of decision trees. Each tree is used to do prediction about the packet as normal or attack. Random forest ensembles the output coming from all trees and generated the final output based on majority voting concept. All the trees are trained with same dataset. Training can be done using same subset of dataset for all trees it is called as homogeneous training. Other training can be heterogeneous where all trees are trained using different datasets. Training of trees is decided according to complexity of dataset and network traffic. Accuracy of dataset is more when sufficient number of trees is used. This classifier provides highest accuracy for labelled data in supervised machine learning algorithm.

3.4 Proposed System Architecture

System architecture for proposed IDS is shown in Fig 3. In the proposed system IDS is implemented using machine learning algorithms as classifiers. Three classifiers used for development is naïve bayes, decision tree and random forest. These all classifiers are trained using new customized and reduced dataset. All training is done using 25 features which are reduced from existing features of NSL-KDD dataset. As

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

shown in fig. 3 IDS is implemented using three classifiers as random forest, naïve bayes and decision tree. These all classifiers are trained using customized dataset having signatures of single and multi-attacks. All the classifiers are trained using reduced 25 attributed as discussed above. These features have strong relation with each other so selected for attack detection. All the three classifiers provide good accuracy and detection rate with reduced false alarm rate. We are trying to implement signature type of IDS as all the signatures are stored in dataset and classifiers are trained using it. All the predictions of these three classifiers are passed to ensemble model which works on majority voting algorithm. This algorithm takes input from all three classifiers and takes final decision according to majority votes for attack or normal.

If the packet is predicted as attack by ensemble model then alert is generated by the system and informed to the administrator. If prediction is normal then system allows packet to enter in the network. IDS work efficiently and provides very good detection rate for both single and multi-attack. False alarm rate is less for this system. Performance of IDS using feature selection and ensemble is compared with individual classifier. Ensembling approach is used to improve performance of IDS. Ensembling method used is bagging which makes use of majority voting algorithm to give final prediction about the traffic. Majority voting algorithm works on vote's method. This algorithm collects predictions from each classifier which is consider as vote and depend on majority final prediction is made. For example if out of three classifiers, two classifiers predict traffic as attack then final prediction will be attack. This algorithm is used in bagging fashion to improve performance of weak classifiers. Results and discussion are explained in next section.

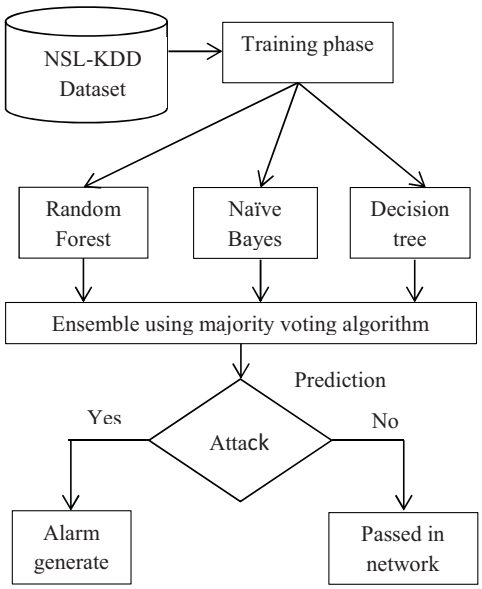


Fig. 3 System Architecture of Proposed System

IV. RESULT AND DISCUSSIONS

All three classifiers are tested for single and multi-attack detection with help of new customized dataset. Performance of the classifiers is evaluated using precision, accuracy and false alarm rate parameters. All classifiers are tested individually on dataset along with ensemble technique. Parameters used to test performance of IDS follows confusion matrix. Confusion matrix is a performance measure for evaluation of performance of IDS. When classification of data is done to check its efficiency and accuracy performance of IDS is checked with the help of confusion matrix. It is matrix for predicted and actual values used to check performance of any prediction done by IDS. Matrix used for evaluation is as shown below in Fig. 4

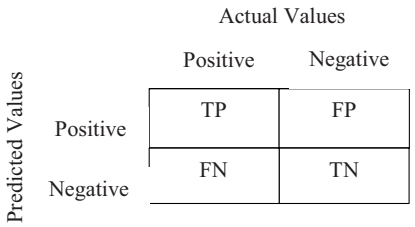


Fig. 4 Confusion matrix

As shown in matrix predicted values and actual values are used to evaluate the performance of IDS. These values are used to calculate accuracy, precision and many more parameters.

TP – Prediction is positive value and the prediction is correct.

FP – Prediction is positive value and the prediction is incorrect

TN – Prediction is negative value and the prediction is correct

FN – Prediction is negative value and the prediction is incorrect

All these values are used to calculate precision and accuracy. Formulas for calculating it is given in Equation (1) and (2) respectively

$$precision = \frac{TP}{TP + FP} \tag{1}$$

$$accuracy = \frac{TP + TN}{total} \tag{2}$$

Where total – total number actual values

Another parameter false positive rate is also used for evaluation. FAR is the total number of wrong predictions made out of actual value. All the classifiers are trained using updated NSL-KDD dataset which consist of signatures of single as well as multi-attack. Each classifier performance is

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

tested using detection rate over all classes of NSL-KDD dataset. Dataset consist of four categories of attacks as probe, DoS, u2r and r2l. One more category is added after adding signatures of multi-attack. Detection rate for each classifier is analysed separately for presenting overall performance of all classifiers including ensemble approach. Table 1-4 shows detection rate of each classifier over all classes of NSL-KDD dataset.

Table 1 Performance evaluation of DT classifier over all classes of NSL-KDD

Attack Classes	normal	probe	Dos	u2r	r2l	Detection rate
Probe	3871	1530	1220	350	713	50.37%
Dos	653	15835	14835	21	178	50.24%
u2r	655	1863	1963	10	12	43.60%
r2l	257	454	614	123	15	8.40%
Multi-attack	169	510	460	56	42	3.39%

Table 2 Performance evaluation of NB classifier over all classes of NSL-KDD

Attack Classes	normal	probe	DoS	u2r	r2l	Detection rate
Probe	3271	1130	1020	150	513	53.77%
Dos	453	17935	11835	41	138	59%
u2r	555	1363	2163	12	10	52.71%
r2l	157	412	414	423	10	29.87%
Multi-attack	199	410	420	36	22	2.06%

Table 3 Performance evaluation of RF classifier over all classes of NSL-KDD

Attack Classes	normal	probe	DoS	u2r	r2l	Detection rate
Probe	2271	1530	1220	110	313	41.71%
Dos	353	12935	1835	71	188	84.09%
u2r	155	1163	2063	11	10	60.64%
r2l	137	212	314	323	9	32.46%
Multi-attack	189	310	350	26	12	1.35%

Table 4 Performance evaluation of Ensemble classifier over all classes of NSL-KDD

Attack Classes	normal	probe	DoS	u2r	r2l	Detection rate
Probe	4771	1230	1120	100	113	65.05%
Dos	343	12935	1535	91	168	85.82%
u2r	145	1063	2663	11	10	68.42%
r2l	187	112	318	383	10	37.92%
Multi-attack	179	110	320	56	99	12.95%

Graphical representation for all classes of attacks is represented in Fig.5.

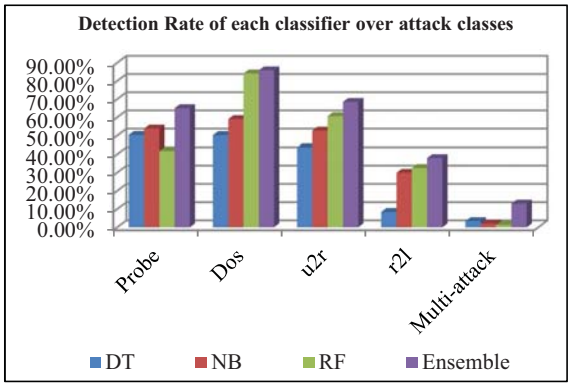


Fig. 5 Comparison of detection rate of each classifier

From Fig. 5 we can observe that detection rate of ensemble approach is highest among individual classifier. Figure shows comparison over all classes of attack signatures available in NSL-KDD dataset. All performance parameters are tested for classifiers individually and in ensemble approach. Comparison of it is shown in table 5 below.

Table 5 Experiment Results for Classifiers

Classifier	Precision	FAR	Accuracy
Naïve bayes	0.89	0.17	89.55%
Decision tree	0.92	0.18	90.22%
Random Forest	0.98	0.13	96.79%
Ensemble approach	0.99	0.09	98.76%

As per observations from table 5 we can conclude that ensemble approach shows better performance in detection of single attack and multi-attack. Graphical representation for all parameters with classifier is shown in Fig. 6 below. Figure shows reduction in false alarm rate for ensemble approach and improvement in precision and accuracy of IDS.

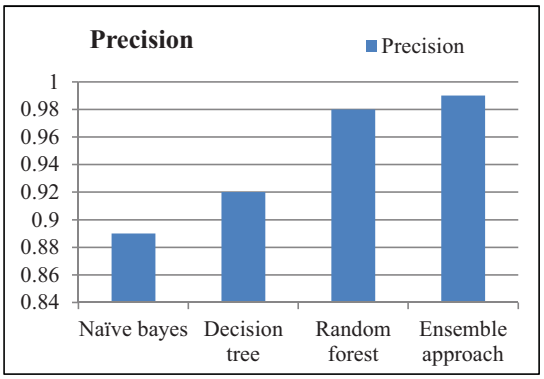


Fig. 6 (a) Graphical representation of precision

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

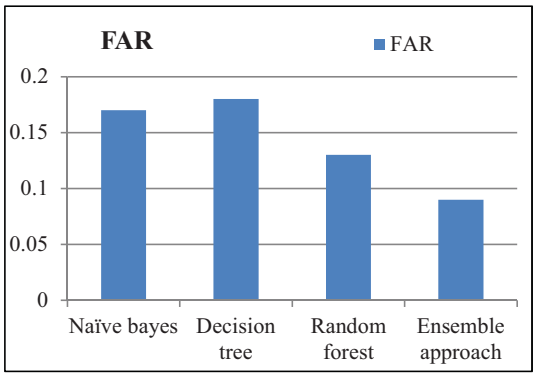


Fig. 6 (b) Graphical Representation of False alarm rate

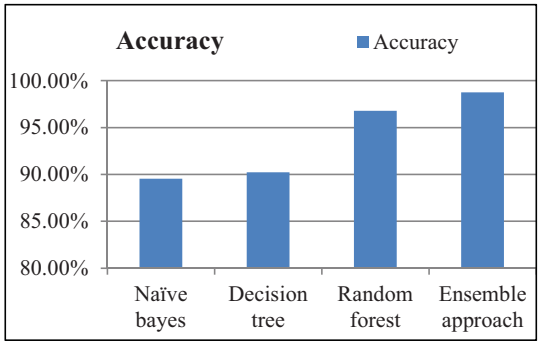


Fig. 6 (c) Graphical Representation of Accuracy

Parameters precision and accuracy which shows the strength of IDS is more when used in ensemble approach. FAR is low means it does correct predictions for most of the records and packets. This approach can be used for anomaly based IDS as well to detect novel multi-attacks. Results show 0.05% reduction in false alarm rate and 2% improvement in accuracy of IDS. Precision of ensemble is also increased by 0.01. This proves that performance of IDS using single classifiers is poor as compared to performance of IDS using ensemble approach. Ensemble approach removes prediction incorrectness and limitations of biased predication.

V. CONCLUSION

The approach presented in paper provides better efficient and robust intrusion detection system. As none of the system available now days does detection of multi-attack in the network. This system allows us to do single as well as multi-attack detection. To detect multi-attack signatures of it are created with the help of scripts. Scripts are executed and multi-attack traffic was generated to penetrate in the network. Traffic with normal packets and attack packets are captured into a pcap file. After pre-processing feature extractor is used to create signature of multi-attack. These signatures are saved into the dataset with other available signatures. Base dataset used is NSL-KDD. Customization of this dataset is done to merger signatures of single and multi-attacks. Feature selection technique is used to reduce number of features of the dataset. Machine learning classifiers decision tree, naïve bayes and random forest are trained for detection of attacks To provide robust and efficient output from IDS ensemble approach is used. Comparison of individual classifier with ensemble approach is presented. Results prove increase in precision and accuracy of detection with less false alarm rate. From results we can conclude that IDS shows good and efficient performance when implemented using ensemble approach and feature selection.

REFERENCES

[1] M. A. Jabbar, Rajanikanth Aluvalu, Sai Satyanarayana Reddy S, "RFAODE: A Novel Ensemble Intrusion Detection System", Procedia Computer Science, vol. 115, 2017, pp. 226-234. DOI: 10.1016/j.procs.2017.09.129

[2] Divyasree T.H, Sherly K.K," A Network Using Intrusion Detection System Based Approach On Ensemble CVM Efficient Feature Selection Using Efficient Feature Selection Approach", Procedia Computer Science, vol. 143, 2018, pp. 442-449. DOI: 10.1016/j.procs.2018.10.416

[3] Kunal, Mohit Dua," Attribute Selection and Ensemble Classifier based Novel Approach to Intrusion Detection System", Procedia Computer Science, Vol. 167, 2020, pp. 2191-2199. DOI: 10.1016/j.procs.2020.03.271

[4] Vinutha, H. P, and B. Poornima," An ensemble classifier approach on different feature selection methods for intrusion detection", Information systems design and intelligent applications, vol. 672, 2018, pp. 442–451.

[5] Govind Gupta and Manish Kulariya, "A Framework for Fast and Efficient Cyber Security Network Intrusion Detection Using Apache Spark", Procedia Computer Science, vol. 93, 2016, pp. 824–831. https://doi.org/10.1016/j.procs.2016.07.238

[6] Gianluigi Folino, Pietro Sabatino,"Ensemble ssbased collaborative and distributed intrusion detection systems: A survey", Journal of Network and Computer Applications, vol. 66, 2016, pp. 1-16. https://doi.org/10.1016/j.jnca.2016.03.011

[7] Gaikwad, D. P, and Ravindra C. Thool, "Intrusion detection system using Bagging with Partial Decision Tree base classifier." Procedia Computer Science Vol. 49, 2015, pp. 92–98. https://doi.org/10.1016/j.procs.2015.04.231

[8] Faezah Hamad Almasoudy, Wathiq Laftah Al-Yaseen , Ali Kadhum Idrees," Differential Evolution Wrapper Feature Selection for Intrusion Detection System", Procedia Computer Science, vol. 167, 2020, pp. 1230-1239. DOI: 10.1016/j.procs.2020.03.438

[9] T.Saranya, S.Sridevi, C.Deisyc, Tran Duc Chungd, M.K.A.Ahamed Khane," Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review", Procedia Computer Science, vol. 171, 2020, pp. 1251-1260. DOI: 10.1016/j.procs.2020.04.133

[10] Peiying Tao, Zhe Sun, Zhixin Sun," An improved intrusion detection algorithm based on GA and SVM", IEEE Access, vol. 6, 2018, pp. 13624-13631. DOI:10.1109/ACCESS.2018.2810198

[11] Abdulla Amin Aburomman, Mamun Bin Ibne Reaz,"A survey of intrusion detection systems based on ensemble and hybrid classifiers", Computer and Security, vol. 65, 2017, pp. 135-152. https://doi.org/10.1016/j.cose.2016.11.004

[12] Salman Niksefat, Parisa Kaghazgaran, Babak Sadeghiyan,"Privacy issues in intersion detection systems: A taxonomy, survey and future directions", Computer Science Review, vol. 25, 2017, pp. 69-78. https://doi.org/10.1016/j.cosrev.2017.07.001

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

[13] Neminath Hubballi, Vinoth Suryanarayanan," False alarm minimization techniques in signature-based intrusion detection systems: A survey", Computer Communications, vol. 49, 2014, pp. 1-17. https://doi.org/10.1016/j.comcom.2014.04.012

[14] R. Vijayanand, D. Devaraj, B. Kannapiran," Intrusion detection system for wireless mesh network using multiple support vector machine classifiers with genetic-algorithm-based feature selection", Computers and Security, vol. 77, 2018, pp. 304-314. https://doi.org/10.1016/j.cose.2018.04.010

[15] Muhammad Fahad Umer, Muhammad Sher, Yaxin Bi," Flow-based intrusion detection: Techniques and challenges", Computers and Security, vol. 70, 2017, pp. 238-254. https://doi.org/10.1016/j.cose.2017.05.009

[16] Jie Gu, Lihong Wang, Huiwen Wang, Shanshan Wang, " A novel approach to intrusion detection using SVM ensemble with feature augmentation" , Computers and Security, vol. 86, 2019, pp. 53-62. https://doi.org/10.1016/j.cose.2019.05.022

[17] S.R.Khonde and V. Ulagamuthalvi, "Hybrid Architecture for Intrusion Detection System", Ingenierie des Systemes d'Information, vol. 24, no. 1, 2019, pp. 19-28.

[18] Jinping Liu, Jiezhou He, Wuxia Zhang, Tianyu Ma, Zhaohui Tang, Jean Paul Niyoyita, Weihua Gui, " ANID-SEoKELM: Adaptive network intrusion detection based on selective ensemble of kernel ELMs with random features", Knowledge-based systems, vol. 177, 2019, pp. 104-116. https://doi.org/10.1016/j.knosys.2019.04.008

[19] Srilatha Chebrolu, Ajith Abraham, Johnson P. Thomas," Feature deduction and ensemble design of intrusion detection systems", Computers and Security, Vol. 24, no. 4, 2005, pp. 295-307. https://doi.org/10.1016/j.cose.2004.09.008

[20] Shadi Aljawarneh, Monther Aldwairi, Muneer Bani Yassein, " Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model", Journal of Computational Science, vol. 25, 2018, pp. 152-160. https://doi.org/10.1016/j.jocs.2017.03.006

[21] S.R.Khonde and V. Ulagamuthalvi, "Fusion of Feature Selection and Random Forest for an Anomaly-Based Intrusion Detection System", Journal of Computational and Theoretical Nanoscience, vol. 16, 2019, pp. 1–5.

[22] Yang Li, Jun-Li Wang, Zhi-hong Tian, Tian-Bo Lu, Chen Young," Building lightweight intrusion detection system using wrapper-based feature selection mechanisms", Computers and Security, vol. 28, no. 6, 2009, pp. 466-475. https://doi.org/10.1016/j.cose.2009.01.001

[23] Sara Mohammadi, Hamid Mirvaziri, Mostafa Ghazizadeh-Ahsaee, Hadis Karimipour," Cyber intrusion detection by combined feature selection algorithm", Journal of Information Security and Applications, vol. 44, 2019, pp. 80-88. https://doi.org/10.1016/j.jisa.2018.11.007

[24] Markus Ring, Sarah Wunderlich, Deniz Scheuring, Dieter Landes, Andreas Hotho," A survey of network-based intrusion detection data sets", Computers and Security, Vol. 86, 2019, pp. 147-167. https://doi.org/10.1016/j.cose.2019.06.005

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:54:14 UTC from IEEE Xplore. Restrictions apply.

ScienceDirect

Intermediate Temperature Solid Oxide Fuel Cells
Electrolytes, Electrodes and Interconnects

2020, Pages 113-163

Chapter 5 - Enhancing the ionic conductivity in the ceria-based electrolytes for intermediate temperature solid oxide fuel cells

S. Ajith Kumar, P. Kuppusami

Centre of Excellence for Energy Research, Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, India

Available online 24 January 2020, Version of Record 24 January 2020.

Show less

Outline | Share Cite

https://doi.org/10.1016/B978-0-12-817445-6.00005-3

Get rights and content

Abstract

This chapter explains the emerging trends and development of cerium oxide–based material as a potential solid electrolyte for intermediate temperature solid oxide fuel cells (IT-SOFCs) application. The new strategies and approaches toward the enhancement in the ionic conductivity of the solid electrolyte in the current literature are discussed. The influence of structural and microstructural modifications on electrical properties of the solid electrolyte has been explained from the fundamental material aspect of IT-SOFCs. The various contributing factors for ionic transport such as grain and grain boundary effect in the bulk nanostructures are described. In addition, the importance of the oxide ionic conductivity to reduce the operational temperature of SOFCs is highlighted. The ongoing research in the development and requirements of thin layer of electrolytes using both chemical and physical methods are illustrated. This chapter highlights on the key development of interlayers which can be used in conjunction with doped ceria electrolytes for IT-SOFCs. The work also reviews the recent progress, advancements, and future directions in the ceria-based electrolytes for IT-SOFCs.

Previous Next

Keywords

Intermediate temperature solid oxide fuel cells; ceria electrolytes; co-doping; nanocomposite; ionic conductivity; activation energy; thin-film electrolytes

Recommended articles

Cited by (3)

Investigation of ceria-molten carbonate electrolyte, composite anode and its catalytical effect on various carbon fuels in molten carbonate direct coal/carbon fuel cell

2023, Fuel

Show abstract

Micro-Structural Design of Soft Solid Composite Electrolytes With Enhanced Ionic Conductivity
2022, Journal of Applied Mechanics, Transactions ASME

Research progress of electrolytematerials in solid oxide fuel cells
2021, Gongneng Cailiao/Journal of Functional Materials

Copyright © 2020 Elsevier Inc. All rights reserved.

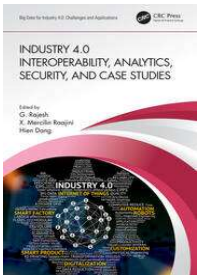


Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

RELX™

Your online security is important to us, so we're taking steps to protect our visitors by ensuring that their browsers always connect to our website over HTTPS. All HTTP requests will no longer be processed/will be blocked. These changes will take place on 27th February 2023.

Chapter



Energy Conservation in IOT-Based Intelligent Transport System

By A. Sivasangari, D. Deepa, Chiung Ching Peter Ho, M.S. Roobini, R. Vignesh

Book Industry 4.0 Interoperability, Analytics, Security, and Case Studies (https://www.taylorfrancis.com/books/mono/10.1201/9781003048855/industry-4-0-interoperability-analytics-security-case-studies?refId=39b6598e-6f9f-46bb-8db4-2b768baffe3c&context=ubx)

Edition	1st Edition
First Published	2021
Imprint	CRC Press
Pages	14
eBook ISBN	9781003048855

Share

ABSTRACT

< Previous Chapter (chapters/edit/10.1201/9781003048855-3/hybrid-huffman-tree-model-securing-medical-data-using-steganography-saranya-mohak-narang-anandan-aruna-balaganesh?context=ubx)
Next Chapter > (chapters/edit/10.1201/9781003048855-5/digital-information-system-attack-reduction-industry-4-0-srinivasan-selvaraj-raja-kothandaraman-kannan-kaliyan-pratheepan-thuraiyappah?context=ubx)

Policies

Journals

Corporate

Help & Contact

Connect with us

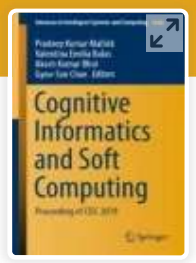
(https://www.linkedin.com/company/taylor-&-francis-group/) (https://twitter.com/tandfnewsroom?lang=en) (https://www.facebook.com/TaylorandFrancisGroup/) (https://www.youtube.com/user/Tayloran

Registered in England & Wales No. 3099067
5 Howick Place | London | SW1P 1WG

© 2023 Informa UK Limited

Springer Link

Search Log in



Cognitive Informatics and Soft Computing, pp 205–214

Encapsulated Features with Multi-objective Deep Belief Networks for Action Classification

Paul T. Sheeba & S. Murugan

Conference paper | First Online: 15 January 2020

442 Accesses

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 1040)

Abstract

Human action classification plays a challenging role in the field of robotics and other human–computer interaction systems. It also helps people in crime analysis, security tasks, and human support systems. The main purpose of this work is to design and implement a system to classify human actions in videos using encapsulated features and multi-objective deep belief network. Encapsulated features include space–time interest points, shape, and coverage factor. Initially, frames having actions had been separated from the input videos by means of structural similarity measure. Later, spatiotemporal interest points, shape and coverage

factor are extracted and combined to form encapsulated features. To improve the accuracy in classification, MODBN classifier was designed by combining multi-objective dragonfly algorithm and deep belief network. Datasets such as Weizmann and KTH are used in MODBN classifier to carry the experimentation. Accuracy, sensitivity, and specificity are measured to evaluate the classification network. This proposed classifier with encapsulated features can produce better performance with 99% of accuracy, 97% of sensitivity, and 95% of specificity.

Keywords

Action recognition **SSIM** **STI** **DBN** **DA** **MODBN**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-1451-7_23
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33
Price includes VAT (India)

- ISBN: 978-981-15-1451-7

- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-1450-0

- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Learn about institutional subscriptions

References

1. Blank, M., Gorelick, L., Shechtman, E., Irani, M., Basri, R.: Actions as space-time shapes (2007)

2. Tong, M., Li, M., Bai, H., Ma, L., Zhao, M.: DKD–DAD: a novel framework with discriminative kinematic descriptor and deep attention-pooled descriptor for action recognition. Neural Comput. Appl. **1** (2019)

3. Jia, C.C., et al.: Incremental multi-linear discriminant analysis using canonical correlations for action recognition. Neurocomputing **83**, 56–63 (2012)

4. Dollar, P., Rabaud, V., Cottrell, G., Belongie, S.: Behavior Recognition via Sparse Spatio-Temporal Features, pp. 65–72. IEEE (2005)

5. Schuldt, C., Barbara, L., Stockholm, S.: Recognizing human actions: a local SVM approach. In: Proceedings of 17th International Conference, vol. 3, pp. 32–36 (2004)

6. Moussa, M.M., Hemayed, E.E., El Nemr, H.A., Fayek, M.B.: Human action recognition utilizing variations in skeleton dimensions. Arab. J. Sci. Eng. **43**, 597–610 (2018)

7. Huynh-The, T., Le, B.V., Lee, S., Yoon, Y.: Interactive activity recognition using pose-based spatio–temporal relation features and four-level Pachinko Allocation model. Inf. Sci. (NY) **369**, 317–333 (2016)

8. Kong, Y., Jia, Y.: A hierarchical model for human interaction recognition. In: Proceedings of IEEE International Conference Multimedia Expo, pp. 1–6 (2012)

9. Bregonzio, M., Gong, S., Xiang, T.: Recognising action as clouds of space-time interest points. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1948–1955 (2009)

10. Liu, J., Shah, M.: Learning human actions via information maximization. In: 26th IEEE Conference Computer Vision and Pattern Recognition, CVPR (2008)

11. Farhadi, A., Endres, I., Hoiem, D., Forsyth, D.: Describing objects by their attributes. In: CVPR 2009. IEEE Conference, pp. 1778–1785 (2009)

12. Wu, D., Shao, L.: Silhouette analysis-based action recognition via exploiting human poses. IEEE Trans. Circuits Syst. Video Technol. **23**, 236–243 (2013)

13. Rodriguez, M., Orrite, C., Medrano, C., Makris, D.: A time flexible kernel framework for video-based activity recognition. Image Vis. Comput. **48–49**, 26–36 (2016)

14. Li, H., Chen, J., Hu, R.: Multiple feature fusion in convolutional neural networks for action recognition. Wuhan Univ. J. Nat. Sci. **22**, 73–78 (2017)

15. Wang, H., Yuan, C., Hu, W., Ling, H., Yang, W., Sun, C.: Action recognition using nonnegative action component representation and sparse basis selection. IEEE Trans. Image Process. **23**(2), 570–581 (2014)

16. Li, W.X., Vasconcelos, N.: Complex activity recognition via attribute dynamics. Int. J. Comput. Vis. **122**, 334–370 (2017)

17. Nigam, S., Khare, A.: Integration of moment invariants and uniform local binary patterns for human activity recognition in video sequences. Multimed. Tools Appl. **75**, 17303–17332 (2016)

18. Hasan, M., Roy-Chowdhury, A.K.: A continuous learning framework for activity recognition using deep hybrid feature models. IEEE Trans. Multimed. **17**, 1909–1922 (2015)

19. Meng, H., Pears, N., Bailey, C.: Human action classification using SVM_2K classifier on motion features, pp. 458–465 (2006)

20. Everts, I., Van Gemert, J.C., Gevers, T.: Evaluation of color spatio-temporal interest points for human action recognition. IEEE Trans. Image Process. **23**, 1569–1580 (2014)

21. Laptev, I., Lindeberg, T.: Velocity adaptation of space-time interest points. In: Proceedings of International Conference on Pattern Recognition, vol. 1, pp. 52–56 (2004)

22. Vojt, J.: Deep neural networks and their implementation (2016)

23. KTH dataset from, http://www.nada.kth.se/cvap/actions/

24. Weizmann dataset from, http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html

25. Sopharak, A., Uyyanonvara, B., Barman, S., Williamson, T.H.: Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods. Comput. Med. Imaging Graph. **32**, 720–727 (2008)

Author information

Authors and Affiliations

Faculty of Computer Science & Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Paul T. Sheeba

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

S. Murugan

Corresponding author

Correspondence to Paul T. Sheeba .

Editor information

Editors and Affiliations

School of Computer Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University, Bhubaneswar, Odisha, India

Dr. Pradeep Kumar Mallick

Faculty of Engineering, Aurel Vlaicu University of Arad, Arad, Romania

Prof. Dr. Valentina Emilia Balas

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Rangpo, India

Prof. Akash Kumar Bhoi

Division of Information and Communication, Baekseok University, Cheonan-si, Ch'ungch'ong-namdo, Korea (Republic of)

Dr. Gyoo-Soo Chae

Rights and permissions

Reprints and Permissions

Copyright information

© 2020 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Sheeba, P.T., Murugan, S. (2020). Encapsulated Features with Multi-objective Deep Belief Networks for Action Classification. In: Mallick, P., Balas, V., Bhoi, A., Chae, GS. (eds) Cognitive Informatics and Soft Computing. Advances in Intelligent Systems and Computing, vol 1040. Springer, Singapore. https://doi.org/10.1007/978-981-15-1451-7_23

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-1451-7_23

Published	Publisher Name	Print ISBN
15 January 2020	Springer, Singapore	978-981-15-1450-0

Online ISBN	eBook Packages
978-981-15-1451-7	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

9th IEEE International Conference on Communication Systems and Network Technologies

Emotion Classification on Twitter Data Using Word Embedding and Lexicon Based Approach

R.Jeberson Retna Raj, Prasanjeet Das, Prabat Sahu
Department of Information Technology, School of Computing,
Sathyabama Institute of Science and Technology,
Chennai, India.
Jebersonretnarajr@gmail.com

***Abstract*—Twitter is one of the leading social media platforms for its users for tweeting 280 characters in a single tweet. Social media influence the users to share data, promoting advertisement and posting useful information to the followers. The social media data helps the individuals and business personals to take decision based on the analysis with the data. Twitter sentimental analysis is important for identifying the similar text patterns present in the given input text. The analysis further classified with emotional, positive and negative tweets. Existing approaches limited in accuracy, therefore, the word embedding and lexicon based approach has been introduced for increasing the accuracy. The twitter data stream is taken as an input and preprocessed by removing stop words, hash tags and urls. Further, the data is tokenized and applied with the word embedding method to detect the location and lexicon based approach to segregate the sentimental and emotional tweets. The system has been tested with the live data set as well as offline dataset, and the result shows very promising.**

Keywords—Sentiment Analysis; Wordmodel; Classification;

I. INTRODUCTION

Social media plays vital role in reaching out its users for communicating, sharing and socially engaged with others. The survey shows that approximately 2.62 billion social media users actively worldwide and it is keep on growing. Social media widely used for sharing information, knowledge and many different ideas across communities organization and so forth. Micro blogs are important due to its shorter version of media data that can be entered and viewed by its users. Nowadays, data is an important asset. Social media is a source of information which plays critical role in companies for decision making, administrative management and political campaign etc. The micro blogging site Twitter is a powerful medium which is used by millions of users in social networks. Therefore, the conveyed message reaches the users immediately. Crowd sensing is important to sense the twitter users for analyzing the tweets with emotional and sentimental pattern. Furthermore, location based crowd sourcing is famous for identifying location based on tweet data. Emotional mining used to extract the relevant emotional patterns from the tweets and categorize them. The micro blogs data is composed of comments, reviews and discussions with the combined data of textual, image and emoji features. In essence, these data are noisy in nature so that it should be processed such a way that to make it as noise free. The community users express their thoughts through casual or formal approaches. Due to these methods, noisy data are incorporated which leads to lack of conformance across communities and individuals. The natural language processing algorithms doesn't handle these kinds of difficulties. Artificial Intelligence enabled Natural Language Processing (NLP) techniques processes the human language and transform into world understanding format. Furthermore, NLP is a foundation for many of the text processing techniques. Generally, the data analytics algorithm faces the challenge as incompleteness of data. For example, microblogging applications have the provision that the user can disable the Geo-location. This further leads ineffective for most of the crowd source sensing techniques.

A. Crowd Source Sensing

Mobile crowd-sourced sensing is emerged as a new sensing paradigm which is a user-mobile centric method for collecting data from mobile users and devices. In this technique, smart phones are act as a physical sensor which collects the data as volunteered source. Smart phone users are the source of required information which make use of images and related data. Through smart phones Collecting data from mass applications like Swiggy, Zomato, Ola, Uber, Zoomcar etc., for further analysis. In the proposed work, twitter as a crowd-source for getting data and appropriate information, which collects data from real time tweets. With the help of geo-location as parameter for filtering the tweets, we can easily identified whether the tweets from India or other country.

B. Emotion Classification Review

Computer science researchers frequently used the terms like emotion, feelings and mood in emotion mining problems. Emotion is defined as personal and relative reaction to inner or outer actions having consequences of the organism specifically. It is for short time period. The term feeling defined as the personal explanation of emotion, private in experiencing them specifically which exists for very small duration. The term mood defined as sentimental state of mind which is related to emotion exists for long duration but less concentrated.

978-1-7281-4976-9/20/$31.00 ©2020 IEEE
DOI: 10.1109/CSNT.2020.29

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 17:02:44 UTC from IEEE Xplore. Restrictions apply.

C. Problem statement

The current technology used for crowd-source sensing is based on social data experiences many shortcomings which are not addressed. Identifying geo-location of user is very difficult as the user does not provide the geographical coordinates of user's activities. The existing classification techniques giving less importance in preprocessing to filter noise which plays major role in increasing the accuracy of the algorithm. Furthermore, the existing works are limited in the field of crowd-sensing which requires attention. In this work, real time Twitter data is utilized by the framework to address the above problems.

The contribution of the work is divided into three folds. First, the tweets are fetched through Twitter API and with the help of prepreocessing method the url, hashtags and the sysmbol @ are removed and cleaned the data. Second, with the help of word embedding approach the location of tweets can be identified. Finally, with the help of lexicon based approach the sentiment and emotional tweets are classified.

II. Literature Survey

Numerous works has been done in sentimental analysis with the Twitter data. Jianqiang et al presented a deep CNN model for sentimental analysis with the twitter data. The method making use of word embedding method to correlate the semantic relationship and co-occurrences of statistical features between tweets are analyzed. The sentiment feature of word is obtained through the combined features of word embedding and n-gram feature with sentiment polarity score is taken [1]. Saad et al presented a study on twitter sentimental analysis using regression algorithm. Machine learning algorithms are applied with the twitter data set to classify the sentimental and emotional details. The tweets of positive tweets and negative tweets are categorized. They have undergone the process of removal of hastags(#,@), remove stop word, tokenization, stemming, case conversion, feature extraction balancing and scoring and finally classification[2]. Bibi et al proposed an imputation algorithm CIARAA applied to twitter data set to improve the accuracy of classification result. Preprocessing is an important step to fine tune the performance of the classifier. Dimensionality reduction mechanism is used to reduce the input data which increase the accuracy. Furthermore Naïve bayers and SVM algorithm is used to classify the tweets [3]. Feizollah et al presented a sentimental analysis over halal products in twitter. The deep learning algorithms CNN, LSTM and RNN algorithms are used to apply with the twitter data set with the help of word2vec feature. The tweet of a span of ten years has been taken and the sentiments of user with the halal products are predicted [4]. Jianqiang et al presented a preprocessing model for increase the accuracy of classification result. The n-gram approach and polarity feature model is used to preprocess the text data [5]. Phan et al presented a model to address the fuzzy sentiments in tweet data. The raw tweets are cleaned, preprocessed, feature extracted and sentiment model is used for analysis. Finally, the sentimental polarity model used to determine strong positive, positive, neutral, negative and strong negative tweet can be identified [6]. Bouazizi et al presented a method to classify the sentimental details in the tweet using multi-class classifier algorithms [7]. In [8], AI based methods to correlate the social media with socio-economic recovery activities are analyzed. In[9] , an ensemble CNN model is used to classify the personal medicine intake level of sentimental and emotional tweets in the twitter data. In [10], various social media analytical models are introduced to address the different kind of applications. Xu et al presented a extended dictionary sentimental words for improving accuracy over sentiment analysis. The dictionary contains text, field and poly sentiment words for improve the efficiency of the system. Naïve Bayesian classifier is used to map the text with poly sentiment words [11]. Bouazizi et al presented a SENTA tool used to classify seven different sentiments with the given text data [12].

III. Proposed Method

A. Real time Twitter data

The real time twitter data can be fetched from twitter API using python program. The Twitter API provides multiple functionalities which includes the parameters like requirement of Tweets quantity, time interval for the Tweets are posted and geo-location of the posts. The geo-location identification is very helpful in crowd-source sensing, but all users does not enables their geographical coordinate (geo-location) function in the application.

B. Word model and crowd-source sensing

Crowd-source sensing from twitter data discussed earlier is not applicable for several reasons, which includes the fact that all users does not share information about their geo-location. This problem is solved with the help of word model technique which allows getting the information about the user's geo-location [13]. The pseudo code for the word model technique is given below:

```
Input: raw tweets from twitter data (RT)
Output: filtered tweets (FT)
Parameters: location, city, place
Begin
    For all I to n do
    Begin
        If RT[i]==location
            FT=RT
        endif
    End
    For all i to n do
    Begin
        If RT[i]==city then
            FT=RT
          if RT[i]==place then
            FT= RT
           else
            discard
           endif
        endif
    end
```


Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 17:02:44 UTC from IEEE Xplore. Restrictions apply.

The above algorithm states that if the twitter account users has disabled their geo-location function, with the help of this algorithm the geo-location of the user can be identified which can be the name of city, country or any venue. The algorithm takes the real time Tweets as a data input fetched from Twitter. Tweet goes directly to the cleaning stage if the geo-location function is enabled or available. If other cases exists like if the geo-location is not enabled then the data goes through the above algorithm processing in order to get the location information of the tweets. It is not always possible for the user to provide location information.

C. *Data Preprocessing*

In tweets, there is a possibility of casual words with spelling mistakes, special characters, short forms and abbreviations are written by the general public. The input tweets containing these kind of noise which degrades the classification accuracy. Therefore, it should be cleaned by suitable preprocessing algorithm and classifying it based on lexicon based approaches. Data pre-processing methodology are shown in figure below:

- Special characters present in the data will be removed.
- Hyperlinks present in the twitter data will be removed.
- Useless punctuations from the data are removed.
- Numbers, stopping words present in the tweets will be removed.
- All the letters in the uppercase will be converted into lowercase format.

D. *System Architecture*

The architecture of the proposed system is shown in Figure 1. The raw input tweets are collected through Twitter API and stored it in a database. The data preprocessing is used to clean the data by removing hastag, url and @ symbols.

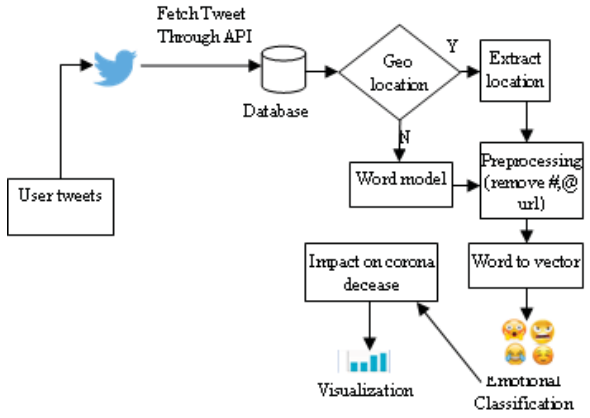


Figure : 1. System architecture of proposed system

The wordmodel algorithm used to segregate the place, Geo-location and city from the tweet data and stored in the database. The lexicon based approach is used to classify the emotional and sentiment over tweet data.

E. *Working procedure of the proposed system*

The steps for the processing of the data and extracting useful data are as follows:

- Collecting data from the twitter using Twitter API.
- Converting useful data to CSV file.
- Dropping the duplicate rows and useless columns.
- Cleaning the Tweets by using word model to clean the noisy data.
- Stopping words removed from the real time data.
- Lemmatizing and Stemming- stemming reduces the words to their stems, lemmatizer does the same work as stemmer but keeping the linguistics of the word context.
- Data Visualisation- It uses word cloud to represent majority of words in each of the sentiments.

F. *Lexicon based classification*

The lexicon based classification algorithm is inputting the preprocessed tweets and output will be classified emotions in text format. The pseudo code of the algorithm is given below:

```
Input : word from lexicon (w)
        Counter variable i
        A sample set (S)
        wj- Occurrence of a word in the text.
Output: classified emotion in text (T)
begin
    For all i in n do
    begin
            If W[i]==S[i] then

                    i=1
            else
                    i=0
            endif
    endfor
    for all i to n do
            i=i*wj
    endfor
    for all i to n do
            return sum of all (T*i) belongs to the set S
    endfor
end
```

IV. EXPERIMENTAL RESULTS

A. *Swatch bharat Review using Twitter Data*

The system is tested with the experimental tweets from Twitter API. For first experiment, a scenario is considered

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 17:02:44 UTC from IEEE Xplore. Restrictions apply.

in which viewer's responses to the government event Swachh Bharat Abhiyan in positive and negative Tweets through analysis. Figure 2 shows Positive and figure 3 shows negative tweets respectively. The responses of the viewer's are framed as a word model below.



Figure 3: Negative responses for swatch bharat

B. Corona virus review using twitter data

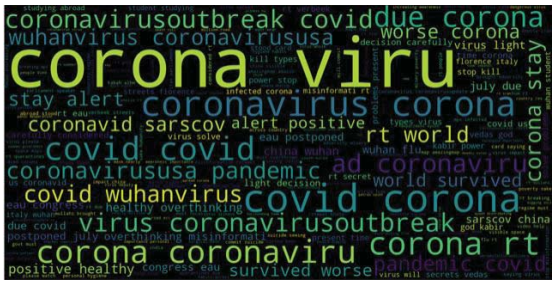


Figure 4: Positive responses for coron virus

The above figure 4 shows the positive reviews on tweets in the twitter about the CORONA VIRUS. Word clouding technique is used to represent the highlighted words in the word cloud.

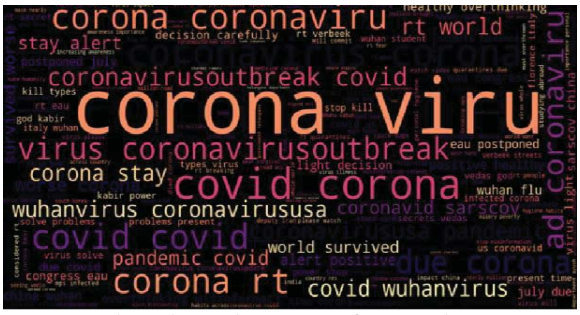


Figure 5: Negative responses for coron virus

The above figure shows the negative reviews on tweets in the twitter about CORONA VIRUS.

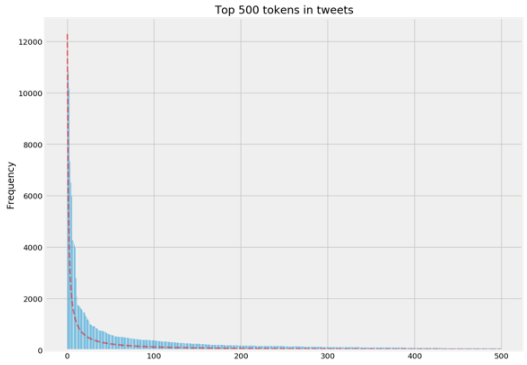


Figure 6: Top 500 tokens in tweets

On the X-axis is the rank of the frequency from highest rank from left up to 500th rank to the right. Y-axis is the frequency observed in the corpus (in this case, "Sentiment140" dataset). One thing to note is that the actual observations in most cases does not strictly follow Zipf's distribution, but rather follow a trend of "near-Zipfian" distribution.

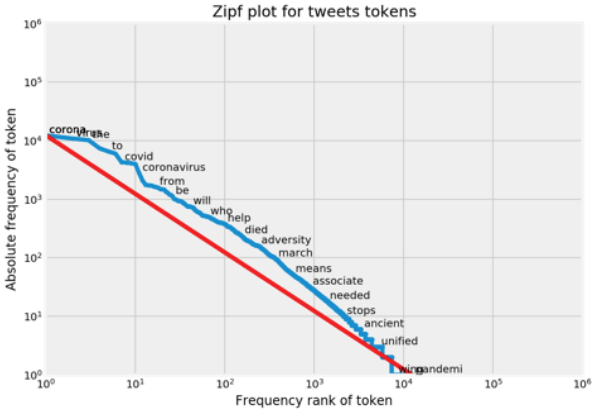


Figure 7: Zipf plot for tweets tokens

Another way to plot this is on a log-log graph, with X-axis being log(rank), Y-axis being log(frequency). By plotting on a log-log scale the result will yield roughly linear line on the graph. Again we see a roughly linear curve, but deviating above the expected line on higher ranked words, and at the lower ranks we see the actual observation line lies below the expected linear line.

Even though some of the top 50 tokens can provide some information about the negative tweets, some neutral words such as "just", "day", are one of the most frequent tokens. Even though these are the actual high-frequency words, but it is difficult to say that these words are all important words in negative tweets that characterizes the negative class. Again, neutral words like "just", "day", are quite high up in the rank.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 17:02:44 UTC from IEEE Xplore. Restrictions apply.

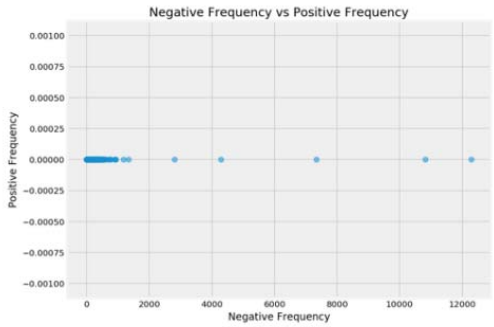


Figure 8: Comparison of positive and negative frequency.

Plot the negative frequency of a word on X-axis, and the positive frequency on Y-axis.

V. CONCLUSION

The system is successfully implemented and tested with the twitter data. Tweepy library is used to collect Tweets which is followed by the selection of useful features in conversion into CSV file. The data then undergoes cleaning process which includes removal of URLs, username-tags, retweet symbols, and hashtags. To identify the sentiments of the tweet SentiWordNet lexicon is used. For mathematical representation of the text count vectorizer and Tfidf is used after splitting the data. For processing the twitter data, steps like lemmatizing, stop words removal stemming are applied. After applying all these data pre-processing steps and applied with the word modeling and lexicon based algorithm is used to classify the tweets. finally the data is visualized using data visualizing technique. This technique visualizes the data in two categories, those are positive responses in tweets and negative responses on the tweet. The system successfully categorizes the sentimental and emotional characteristics of tweets.

REFERENCES

[1] Z. Jianqiang, G. Xiaolin and Z. Xuejun, "Deep Convolution Neural Networks for Twitter Sentiment Analysis," in IEEE Access, vol. 6, pp. 23253-23260, 2018.

[2] S. E. Saad and J. Yang, "Twitter Sentiment Analysis Based on Ordinal Regression," in IEEE Access, vol. 7, pp. 163677-163685, 2019.

[3] M. Bibi et al., "Class Association and Attribute Relevancy Based Imputation Algorithm to Reduce Twitter Data for Optimal Sentiment Analysis," in IEEE Access, vol. 7, pp. 136535-136544, 2019.

[4] A. Feizollah, S. Ainin, N. B. Anuar, N. A. B. Abdullah and M. Hazim, "Halal Products on Twitter: Data Extraction and Sentiment Analysis Using Stack of Deep Learning Algorithms," in IEEE Access, vol. 7, pp. 83354-83362, 2019.

[5] Z. Jianqiang and G. Xiaolin, "Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis," in IEEE Access, vol. 5, pp. 2870-2879, 2017.

[6] H. T. Phan, V. C. Tran, N. T. Nguyen and D. Hwang, "Improving the Performance of Sentiment Analysis of Tweets Containing Fuzzy Sentiment Using the Feature Ensemble Model," in IEEE Access, vol. 8, pp. 14630-14641, 2020.

[7] M. Bouazizi and T. Ohtsuki, "Multi-Class Sentiment Analysis in Twitter: What if Classification is Not the Answer," in IEEE Access, vol. 6, pp. 64486-64502, 2018.

[8] Y. Shibuya and H. Tanaka, "Using Social Media to Detect Socio-Economic Disaster Recovery," in IEEE Intelligent Systems, vol. 34, no. 3, pp. 29-37, 1 May-June 2019.

[9] D. Mahata, J. Friedrichs, R. R. Shah and J. Jiang, "Detecting Personal Intake of Medicine from Twitter," in IEEE Intelligent Systems, vol. 33, no. 4, pp. 87-95, Jul./Aug. 2018.

[10] J. Fiaidhi and S. Mohammed, "Thick Data: A New Qualitative Analytics for Identifying Customer Insights," in IT Professional, vol. 21, no. 3, pp. 4-13, 1 May-June 2019.

[11] G. Xu, Z. Yu, H. Yao, F. Li, Y. Meng and X. Wu, "Chinese Text Sentiment Analysis Based on Extended Sentiment Dictionary," in IEEE Access, vol. 7, pp. 43749-43762, 2019.

[12] M. Bouazizi and T. Ohtsuki, "A Pattern-Based Approach for Multi-Class Sentiment Analysis in Twitter," in IEEE Access, vol. 5, pp. 20617-20639, 2017.

[13] R. Kamal, M. A. Shah, C. Maple, M. Masood, A. Wahid and A. Mehmood, "Emotion Classification and Crowd Source Sensing; A Lexicon Based Approach," in IEEE Access, vol. 7, pp. 27124-27134, 2019.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 17:02:44 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Election Prediction with Automated Speech Emotion Recognition

S.PRAYLA SHYRY
Professor,CSE, Sathyabama Institute of science and technology
Chennai, India
praylashyry@gmail.com

K.KAJA KARTHEEK
Computer science and engineering
Sathyabama institute of science and technology
Chennai, India

K N RR ARAVIND
Computer science and engineering
Sathyabama institute of science and technology
Chennai, India

Abstract- Even though artificial intelligence has made meticulous outcomes, there is a gap and large thirst for speech emotion recognition in affective computing. Since machines don't understand human emotions, natural human-computer interaction is struggling a lot to strive and compete with today's market. The survey of past election results and prediction has become crucial and out-dated since the surveys could never clearly depicts the real mental state of an individual. Since the wealth and familiarity of the political people have a big title role in many surveys, it is not trustworthy. No surveys can understand the truth in the heart of the common people. So a novel approach is developed for day to day physical and psychological behavior of humans which is influenced by emotions. Other than behavioral expressions, emotion recognition from speech signals is a challenging modality. For the implementation, at least 100 audio signals of different areas, village, town, district, community, religion, and privileged people are considered and a nice questionnaire section (20 questions) is used to record the real emotion of the people. MFCC (mel-frequency cepstral coefficient) feature is extracted and classified with a decision tree, random forest, and CNN (convolutional neural network) to give the eight classes of emotions namely neutral, disgust, calm, angry, happy, sad, fearful and surprise. Any traditional algorithm can be compared to determine the performance of speech emotion recognition.

Keywords: Speech Emotion Recognition, MFCC, Decision Tree, Random Forest, CNN.

I. INTRODUCTION

Emotion plays a significant role in understanding and conveying the feelings in day to day life (Male or Female) that are expressed through multiple channels such as linguistic content, voice characteristics, facial expressions, and gestures. There are several ways to recognize human emotions. Some of them are recognition of emotions through speech, facial expressions, and physiological signals. In recent years, there are many applications like entertainment, robots, audio surveillance, web-based E-learning for detecting human emotions. Many algorithms like CNN, SVM (Support Vector Machine), Gaussian Mixture Model(GMM), Hidden Markov Model(HMM) are used to classify different types of emotions like anger, fear, joy, sadness, surprise, etc. by extracting acoustic features like pitch, MFCC, frequency, WPCC(wavelet packet cepstral coefficient).

Emotions can be either positive or negative based on the situation of the person. Researchers believe that speech is the fastest way of communication when compared to other natural methods. Different types of features like MFCC, WPCC are used for extracting spectrograms from the speech signal, utterances from speech signals, and waveforms from speech signals. By using automated speech emotion recognition, several types of research can be performed in the area of emotion recognition, general classification, pattern recognition, etc. which can also be used in the detection of cyber-crimes. Several methods like MFCC extraction, Cost Matrix Method, LPCC (Linear Prediction Cepstral Coefficient) feature extraction methods are used to process speech signals.

II. STATE OF ART

Ismail Shahin et.al (2019) proposed a novel methodology to identify six different emotions using a hybrid cascaded GMM and DNN by using Emirati Arabic - emphasized speech database. They have extracted the MFCC feature from the speech signal which gives the pitch of the human auditory system and logarithmic perception of onset to detect the emotions. For extracting MFCC feature speech signals are converted into discrete cosign transform to acquire spectral coefficients. They have achieved 83.97% by using GMM-DNN classifier while using SVM and MLP they achieved 80.33% and 69.78% to get better system performance in the presence of noisy signals they have provided a dominant signal mask for the hybrid classifier.

Yuan chao li et.al (2019) have classified emotion using spectrogram by extracting features like pitch, MFCC, energy from the IEMOCAP database to perform gender classification. They have used the combination of CNN-BLSTM in the E2E manner for training the model and achieved an absolute increase in overall accuracy by 7.7%bwhencompared to other traditional algorithms. Spectrogram with deltas and delta-deltas are extracted and given as input to the neural network. They have implemented the CNN-BLSTM model using Keras and optimized by Adam method.

Santiago Pascual et.al (2019) have proposed an improved self-supervised method in which multiple workers

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

like waveform, MFCC, LPS prosody solve different self-supervised tasks which are followed by a single neural coder to extract transferable robust problem agnostic features like speaker identity phonemes emotional cues from the speech signal. They discovered robust representation direction from the raw speech waveform using PASE which can be used as a pre-trained network to avoid trained models from scratch.

John W. Kim et.al (2018) have proposed SVM and EMnet model which combines DNN to extract patterns of features with high-level representation and use of common features which is useful for emotion recognition. In this model, they have extracted multi-dimension feature vectors and speech signals for every 10 msec using a 30 msec hamming window. They have used the EMO-DB database which contains 535 speech wave files and 10 short sentences labeled with one of the common emotions. They have achieved 88.9% accuracy from compare across 98 trails.

Jilt Sebastian et.al (2019) have investigated the use of LSTM (long short-term memory) with pre-trained text-based emotion recognition and CNN for emotion recognition with a speech by combining fusion techniques. They have used IEMOCAP multi-model dataset and adopted various fusion strategies to achieve overall score for each emotion. They have extracted features using CNN from utterance transcriptions by taking 300-dimensional pre-trained word embeddings and for each utterance, acoustic features are extracted in each stage for building join CNN model using early fusion to improve performance.

Zheng Lian et.al (2018) have proposed a novel method named pair wise discriminative task (PDT) to obtain high-level audio vectors and compared with traditional classification tasks. Spectrogram based self- attentional CNN-BLSTM model is used to classify emotions in the E2E manner. A total number of 5531 utterances are used to recognize the emotions. They have extracted acoustic and prosodic features like MFCC, pitch, energy using IEMOCAP database and achieved an increase in overall accuracy by 7.7% when compared with other traditional methods.

Wei Jiang et.al (2019) used the DNN algorithm to extract the feature representations from Heterogeneous acoustic feature groups. They have extracted audio data from video data of IEMOCAP dataset to recognize the performance of proposed DNN architecture. To learn discriminative acoustic feature representation to obtain informative features a fusion network is trained jointly. SVM is used to recognize the task and achieved 64% when compared to existing approaches.

Zheng Lian et.al (2019) have combined future observation prediction-unsupervised representation learning strategies with transfer learning approaches to verify the effectiveness of the proposed system. They have conducted experiments to demonstrate the superiority of the proposed system on advanced unsupervised learning strategies by using IEMOCAP database. They have tested each configuration 20 times. They have used Adam optimization scheme with a learning rate of 0.01 and a pack size of 32 to outperform the mainstream named semi-supervised and unsupervised learning strategies above 4% for speech recognition.

Rajeev Ranjan et.al (2019) has filtered audio signal by extracting features like pitch frequency and cepstral coefficients using a feature extraction technique and perform comparative analysis using different types of MFCC feature extraction method. They have transformed audio waveform to frequency-domain representation for advanced processing and analysis. In this proposed system they have analyzed speech in each short frame window to obtain a spectrum by FFT and concluded delta-delta MFCC future extraction technique is better than other feature extraction techniques.

Vladimir Chernykh et.al (2017) have proposed a novel approach named DRNN (deep recurrent neural network) with a sequence of trained acoustic features and compared with recent advances they have classified speech by emotion and time properties of emotions. They have used the IEMOCAP database which consists of 12 hours of audio-visual data. The audio signals are divided into small utterances and each utterance is evaluated from 3 to 4 assessors based on audio and video streams to recognize the emotions. They have compared BLSTM performance in one label approach and CTC approach and predicted the sequence of emotions for one utterance.

Sujay Angadi et.al (2019) has proposed a novel method named CNN-BLSTM to classify human emotions and measure parameters like precision, recall, and accuracy. They have extracted spectral and temporal features to improve SER performance. IEMOCAP database is used to classify emotions by comparing the CNN-BLSTM method with SVM and achieved an increase in 9.83% accuracy.

Hemantha Kumar Palo et.al (2018) have developed a feature extraction technique to reduce combination features by extracting features like wavelet coefficient, LPCC, MFCC for recognizing five different emotions. They have compared the EMO-DB database with the SAVEE database to validate the accuracy and to recognize common arousal emotions by using different ratios of predefined training, testing, and validation data. They have achieved 91.82% and 93.67% accuracy from EMO-DB and SAVEE for the first 16 cepstral coefficients using the RBFN classifier.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

Noushin Hajarolasvadi et.al (2019) has proposed a novel method named k-means clustering to predict the emotion in the speech signals. They have extracted 88 types of different dimensional vectors of audio features like MFCC, Pitch, and Intensity of each frame which is split by the analysis of speech signals. They have achieved 81.05%, 77%, and 72.33% by extracting audio features using SAVEE, RML, and ENTER05 database by applying SVM classifier.

Yea Hoon Kwon et.al (2018) has proposed a CNN model to classify emotion based on multi-model data and improved the accuracy of human emotions. They have combined both EEG (Electro Encephalon Gram) and GSR (Galvanic Skin Response) signals which are pre-processed by Zero crossing rate. They have achieved 73.4% accuracy after noise removed and data processing.

Heekyung Yang et.al (2019) has proposed a novel method multi-column CNN to recognize emotions from Electro Encephalon Gram (EEG) signals. They have used the DEAP dataset and circumplex model with balance and arousal to classify emotions and compare accuracies based on voting and weighted sum and also compared with different traditional algorithms.

III. SUMMARY OF EXISTING WORK

In the existing system, different types of emotions are recognized by using hybrid of cascaded GMM, DNN, CNN-BLSTM, and SVM algorithms by taking different types of datasets like Emirati speech database, IEMOCAP, SAVEE, EMO-DB. They have extracted the most commonly used features like Mel-frequency cepstral coefficient, WPCC, pitch, energy from the speech signal to detect human emotions by taking several trials using trained models. In some model's utterances to classify emotions and to extract the pattern of features and achieved an increase in accuracy for better system performance in the presence of spectrogram, noisy signals, and wave models.

IV. METHODOLOGY

In this project, datasets are collected in audio format from 100 people with different perceptions, religion, and region. Below figure.1 explains the block diagram of the proposed work. In the field visit, different set of questions were posed to the people. A sample questionnaire is given in table 1.

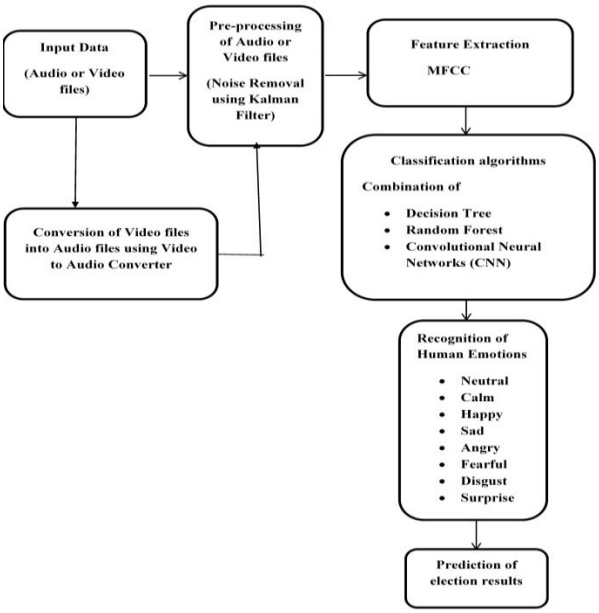


Figure 1. Block Diagram of Speech Emotion Recognition.

Initially, the audio files are pre-processed to remove noise in the audio using Kalman filter and the MFCC feature is extracted. The extracted feature is classified with 3 different classification algorithms and the people with different emotions are evaluated. Also, the accuracy of emotion recognition is calculated to reveal the best classification algorithm.

Table.1:

Sample Questionnaire for the individuals (Takes at least 10 minutes)

1. Do you plan to cast your vote in 2021 Tamil Nadu State election
2. How certain are you that you will vote in the upcoming elections?
3. Which do you identify the most with politically?
4. Do you think the present Government is well enough in decision making.(Justify your answer)
5. Do you think the present Government is well enough in law making. (Justify your answer)

1. Pre-processing:

The .wav file is split into different chunks each of 3 seconds. Each chunk is fed to the Kalman Filter to remove the background noise. Since the background noise (from fan, vehicles) fluctuates over time, it is better to reduce or remove the noise from the wav file to enhance the speech quality. Figure 2 gives the pre-processing of the audio with the Kalman Filter. The outcome of the filter is an enhanced audio signal without any loss of data.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

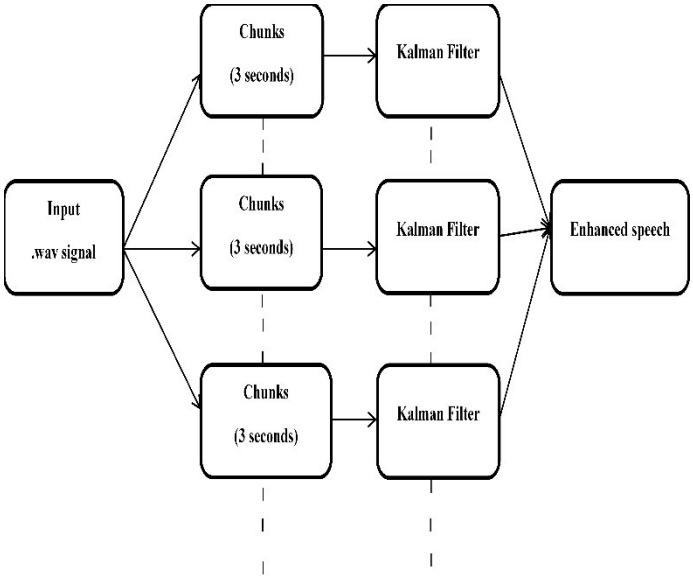
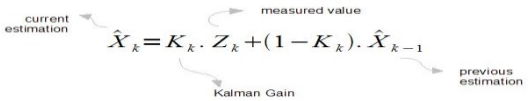


Figure 2. Enhanced Speech

Kalman Filter:

Kalman filter is used to predict unknown values more accurately by taking data inputs from multiple sources. It estimates unknown variables from a high level of signal noise to get an enhanced speech from the wav signal. It is mostly used in navigation and control technology. An Estimation equation $X^{\wedge}_K$ is calculated by adding measured value Z_k and Kalman gain K_K to the previous estimation equation $X^{\wedge}_{K-1}$. Extracting audio features plays a vital role in the classification of emotions.



2. Feature Extraction:

In the proposed system, the MFCC feature has been extracted from the audio signal to identify the human emotion.

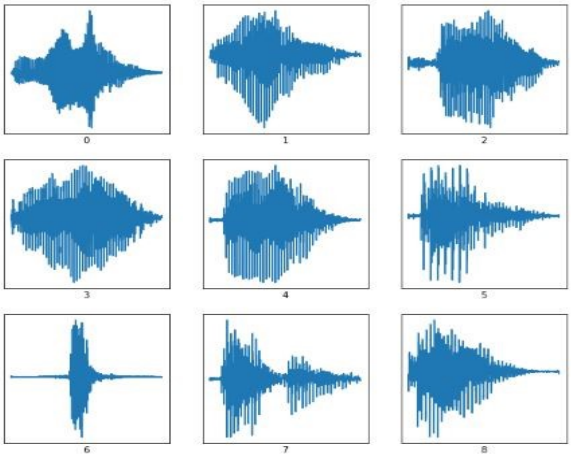


Figure 3. Plotting .wav form of Audio Signal

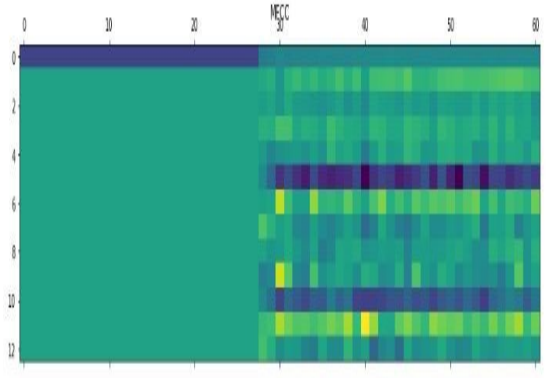


Figure 4. MFCC feature

The coefficients that made up of Mel-Frequency cepstrum are called Mel-Frequency cepstral coefficient. MFCC is derived from a type of cepstral representation of an audio file equally spaced on the Mel scale. In this Fourier transform of the signal is taken and maps the power of spectrum that has been obtained on the Mel scale along with this logs of power and discrete cosine transform is taken and calculated. The amplitude of the resulting spectrum is MFCC.

3. Classification of Emotions:

Emotion recognition helps to understand the real intensity of the person by classifying human emotions. In this proposed system, the audio files are taken as input and pre-processed audio files are used to identify the acoustic MFCC (Mel-frequency Cepstral Coefficient) feature. The extracted feature is classified with a decision tree, random forest, and CNN (Convolutional Neural Network) to give the eight classes of emotions namely Happy, Sad, Neutral, Calm, Disgust, Angry, Fearful and Surprise. Any traditional algorithm can be compared to determine the performance of SER.

3.1 Decision Tree:

A decision tree is easy to understand and used for solving classification and regression problems. The decision tree aims to predict the class or value of the target variable by simple decision terms taken from training data from the created training model,

3.2 Random Forest:

A large number of individual decision trees that operate as an ensemble are called random forest. Class prediction is done in each tree. In the random forest, the class with more predictions is taken as a prediction model.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

3.3 CNN (Convolutional Neural Network):

CNN is a class of deep neural networks mostly used in image processing, speech recognition. CNN consists of Multi-Layer perceptron's (fully-connected networks), each neuron in one layer is connected to all neurons in the next layer. In CNN, Hierarchical patterns are taken from the data and assemble complex patterns using simple and smaller patterns.

V. RESULTS AND DISCUSSION

Precision:

Precision is defined as the fraction of relevant instances among retrieved instances in pattern recognition, informational retrieval, and classification of machine learning algorithms.

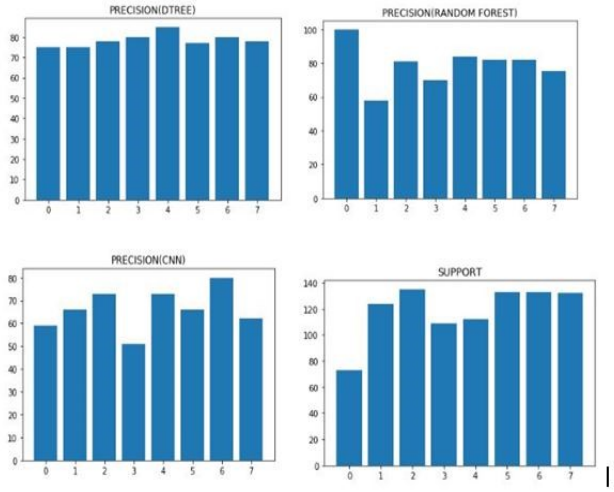


Figure 5. Precision scores

Recall:

A recall is defined as the fraction of the total amount of relevant instances that were retrieved in pattern recognition, informational retrieval, and classification of machine learning algorithms.

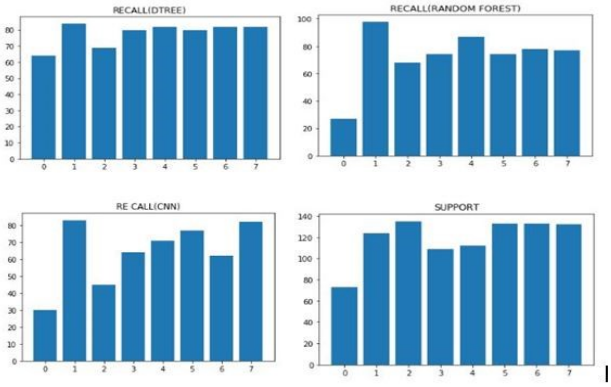


Figure 6. Recall values

F1-Score:

F1 score is defined as the harmonic mean of the precision and recall in the statistical analysis of binary classification.

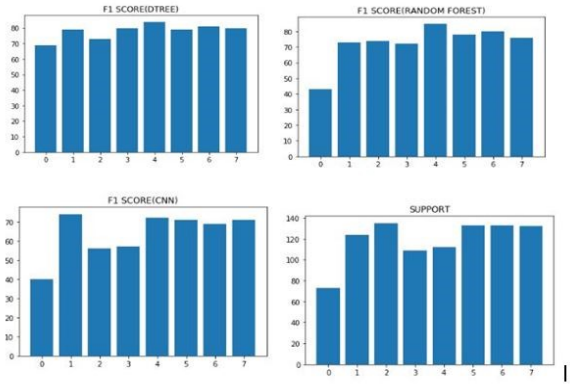


Figure 7. F1 Scores

	precision	recall	f1-score	support
0	1.00	0.27	0.43	73
1	0.58	0.98	0.73	124
2	0.81	0.68	0.74	135
3	0.70	0.74	0.72	109
4	0.84	0.87	0.85	112
5	0.82	0.74	0.78	133
6	0.82	0.78	0.80	133
7	0.75	0.77	0.76	132
accuracy			0.75	951
macro avg	0.79	0.73	0.73	951
weighted avg	0.78	0.75	0.74	951

Decision Tree Confusion matrix

	precision	recall	f1-score	support
0	0.53	0.41	0.46	73
1	0.60	0.81	0.69	124
2	0.69	0.58	0.63	135
3	0.56	0.56	0.56	109
4	0.63	0.74	0.68	112
5	0.64	0.57	0.60	133
6	0.65	0.68	0.66	133
7	0.67	0.65	0.66	132
accuracy			0.63	951
macro avg	0.62	0.62	0.61	951
weighted avg	0.63	0.63	0.62	951

Random Forest Confusion matrix

	precision	recall	f1-score	support
0	0.73	0.67	0.70	73
1	0.76	0.82	0.79	124
2	0.78	0.73	0.76	135
3	0.81	0.80	0.81	109
4	0.79	0.82	0.81	112
5	0.83	0.79	0.81	133
6	0.71	0.79	0.75	133
7	0.81	0.77	0.79	132
accuracy			0.78	951
macro avg	0.78	0.77	0.78	951
weighted avg	0.78	0.78	0.78	951

CNN Confusion matrix

Figure 8. Confusion matrices

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

In the above figure, the confusion matrix of Decision Tree, Random Forest, and CNN are drawn by extracting fl score, precision and recall values

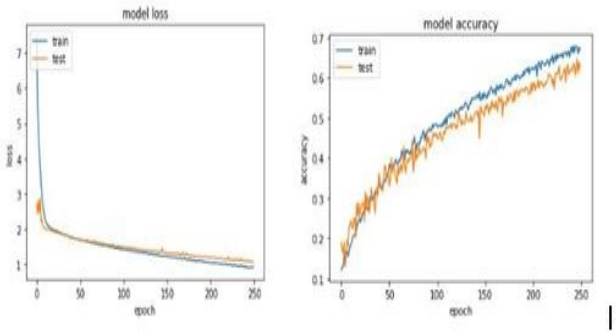


Figure 9. Model loss and Model Accuracy graph

In the above figure model loss and model accuracy are plotted by extracting the loss and accuracy values from the audio signal.

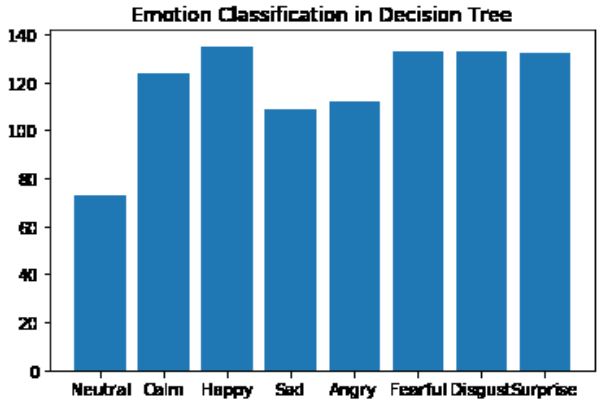


Figure 10. Classification of Emotions using Decision Tree

In the above figure, the emotions extracted from the audio signals using the decision tree are classified and plotted.

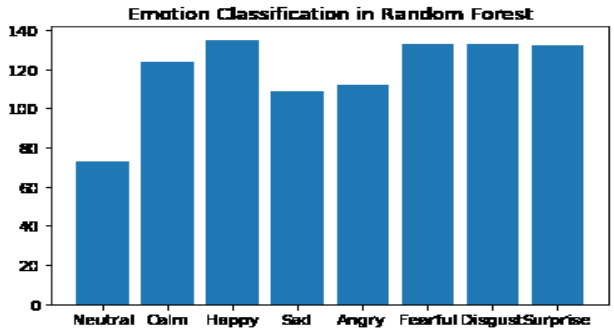


Figure 11.Classification of Emotion using Random Forest

In the above figure, the emotions extracted from the audio signals using random forest are classified and plotted.

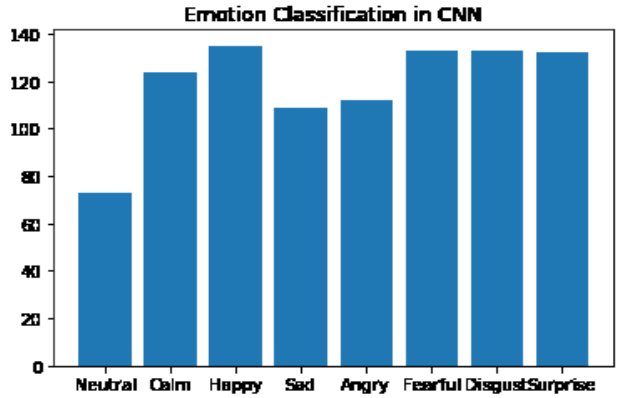


Figure 12. Classification of Emotions using CNN

In the above figure, the emotions that have been extracted from the audio signals using CNN are classified and plotted.

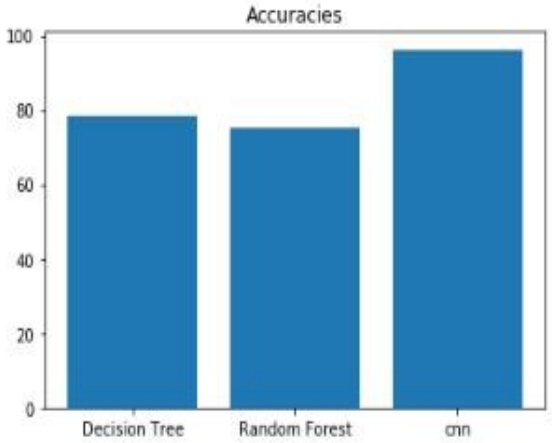


Figure 13. Comparison of accuracy

In the above figure, the accuracies of Decision Tree (79%), Random Forest (75%) and CNN (96%) are compared.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.

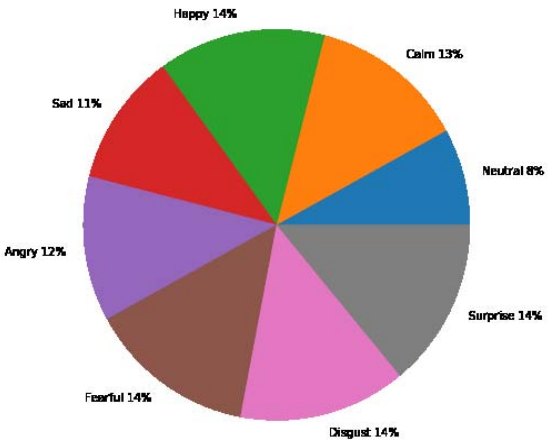


Figure 14. Pie chart of classified Emotions

From the above figure, it has been inferred that 14% of the people are happy, 12% of the people are calm, 8% of the people are neutral, 14% of the people are surprised, 14% of the people are disgust, 14% of the people are fearful, 12% of the people are angry and 11% of the people are sad.

VI. CONCLUSION

This research work has developed a flawless model for speech emotion recognition by extracting acoustic MFCC feature from the audio signals and classify the classes of emotion like neutral, calm, disgust, happy, sad, angry, fearful and surprise using a decision tree, random forest and CNN (Convolutional Neural Network) to achieve an overall increase in accuracy when compared to all traditional algorithms.

REFERENCES:

1. ISMAIL SHAHIN, ALI BOU NASSIF, SHIBANI HAMSA, Emotion Recognition Using. Hybrid Gaussian Mixture Model and Deep NeuralNetwork,http://www.ieee.org/publications_standards/publications/rights/index.html.
2. Jilt Sebastian, Piero Pierucci, Fusion Techniques for Utterance-Level Emotion Recognition Combining Speech and Transcripts, http://dx.doi.org/10.21437/Interspeech.2019-3201.
3. John W. Kim1, Rif A. Saurous, Emotion Recognition from Human Speech Using Temporal Information and Deep Learning, 10.21437/Interspeech.2018-1132.
4. Rajeev Ranjan, Abhishek Thakur, Analysis of Feature Extraction Techniques for Speech Recognition System, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue-7C2, May 2019.
5. Sujay Angadi1* Venkata Siva Reddy1, Hybrid Deep Network Scheme for Emotion Recognition in Speech, International Journal of Intelligent Engineering and Systems, Vol.12, No.3, 2019.
6. Vladimir Chernykh, Grigoriy Sterling, Pavel Prihodko, Emotion Recognition from Speech with Recurrent Neural Networks, arXiv:1701.08071v1 [cs.CL] 27 Jan 2017.
7. Zheng Lian, Jianhua Tao, Bin Liu, Jian Huang, Unsupervised Representation Learning with Future Observation Prediction for Speech Emotion Recognition, arXiv:1910.13806v1 [eess.AS] 24 Oct 2019.
8. Santiago Pascual, Mirco Ravanelli, Joan Serra, Antonio Bonafonte, Yoshua Bengio, Learning Problem-agnostic Speech Representations from Multiple Self-supervised Tasks, arXiv:1904.03416v1 [cs. LG] 6 Apr 2019.
9. Yuanchao Li, Tianyu Zhao, Tatsuya Kawahara, Improved End-to-End Speech Emotion Recognition Using Self Attention Mechanism and Multitask Learning, http://dx.doi.org/10.21437/Interspeech.2019-2594.
10. Yea-Hoon Kwon, Sae-Byuk Shin, and Shin-Dug Kim, Electroencephalography Based Fusion Two-Dimensional (2D)-Convolution Neural Networks (CNN) Model for Emotion Recognition System, www.mdpi.com/journal/sensors.
11. Heekyung Yang, Jongdae Han, Kyungha Min, A Multi-Column CNN Model for Emotion Recognition from EEG Signals, www.mdpi.com/journal/sensors.
12. Noushin Hajarolasvadi, Hasan Demirel, 3D CNN-Based Speech Emotion Recognition Using K-Means Clustering and Spectrograms, www.mdpi.com/journal/entropy.
13. Wei Jiang, Zheng Wang, Jesse S. Jin, Xianfeng Han, Chunguang Li, Speech Emotion Recognition with Heterogeneous Feature Unification of Deep Neural Network, www.mdpi.com/journal/sensors.
14. Zheng Lian, Ya Li, Jianhua Tao, Jian Huang, A PAIRWISE DISCRIMINATIVE TASK FOR SPEECH EMOTION RECOGNITION, www.researchgate.net/publication/322305986_A.
15. Hemanta Kumar Palo, Mihir Narayana Mohant, Wavelet based feature combination for recognition of emotions, https://www.researchgate.net/publication/330263647.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:07:41 UTC from IEEE Xplore. Restrictions apply.



International conference on Computer Networks, Big data and IoT

↳ ICCBI 2018: **Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018)** pp 67–75

Efficiency Improvement in Code Emulation by Soft Computing Technique

S. Prayla Shyry, M. Saranya & M. Mahithaa Sree

Conference paper | First Online: 01 August 2019

1308 Accesses

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 31)

Abstract

Cyber espionage is an illicit way of obtaining secrets or information without user's knowledge. Today, the Stuxnet, Duqu and flame computer worms which are used as cyber weapons against nations continue to make headlines as a new breed of malware. Understanding this topic is important to know how technology shapes the world and influences nation-state relation. It can be in different ways, malicious software like malware- a threat to system security. The technique of malware analysis either in sophistication of samples analyzed through the traditional approaches has its own

con's (Zero day Vulnerability). The aim of this paper is to analyze metamorphic virus using code emulation and implementation of few simple parameters by using fuzzy logic. We also justified how and why these parameters are considered. For a brief introduction to this paper, we used code emulator which creates a virtual environment where we considered number of iterations, memory access function and size of the virus as the performance metrics. The point of code emulation is to mimic the instruction set of CPU using virtual registers and flags.

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-3-030-24643-3_7
- Chapter length: 9 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 128.39
Price includes VAT (India)

- ISBN: 978-3-030-24643-3
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 159.99
Price excludes VAT (India)

- ISBN: 978-3-030-24642-6
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Learn about institutional subscriptions

References

1. Venkatesan, A: Code obfuscation and metamorphic virus detection. Master Thesis, Sanjose state university

2. Szor, P.: The Art of Computer Virus Research and Defense. Addison-Wesley, Boston (2005)

3. Sridharan, S.M.: Hunting for metamorphic engines

4. Wong, W.: Analysis and detection of metamorphic computer viruses

5. Bin, S.A., Sharif, M.D.: Metamorphic worm that carries its own morphing engine

6. Stamp, M., Lin, D.: Hunting for undetectable metamorphic viruses. Master Thesis, Department of Computer science, SJSU, spring 2010

7. Kalbhor, A.: A tiered approach to detect metamorphic malware with hidden markov models

8. Govindaraj, S.: Practical detection of metamorphic virus. Master Thesis, Sanjose state university

9. Pryala Shyry, S.: Performance measurement in selfish overlay network by fuzzy logic deployment of overlay nodes. In: International Conference on Control Instrumentation, Communication Technologies (IEEE Xplore), ICCICCT 2014, pp. 717–721 (2014)

10. Chandrasekharan, M., Muralidharan, M.: Application of soft computing techniques in machining performance prediction and optimization: a literature review. Int. J. Adv. Manuf. Technol. **46**, 445–464 (2010)

11. Choudhary, S.P.: A simple method for detection of metaphoric malware using dynamic analysis and text mining. Procedia Comput. Sci. **54**, 265–270 (2015)

12. Lo, C.T.D., Pablo, O.: Towards an effective and efficient malware detection system. In: International Conference on Big Data (IEEE Xplore)

Author information

Authors and Affiliations

Sathyabama Institute of Science and Technology, Chennai, India

S. Prayla Shyry, M. Saranya & M. Mahithaa Sree

Corresponding authors

Correspondence to S. Prayla Shyry , M. Saranya or M. Mahithaa Sree .

Editor information

Editors and Affiliations

Department of CSE, Vaigai College of Engineering, Melur, India

Prof. Dr. A.Pasumpon Pandian

Department of Engineering, University of the Ryukyus, Okinawa, Japan

Prof. Dr. Tomonobu Senjyu

Edith Cowan University (ECU), Joondalup, WA, Australia

Prof. Syed Mohammed Shamsul Islam

Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

Dr. Haoxiang Wang

Rights and permissions

Reprints and Permissions

Copyright information

© 2020 Springer Nature Switzerland AG

About this paper

Cite this paper

Prayla Shyry, S., Saranya, M., Mahithaa Sree, M. (2020). Efficiency Improvement in Code Emulation by Soft Computing Technique. In: Pandian, A., Senjyu, T., Islam, S., Wang, H. (eds) Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham. https://doi.org/10.1007/978-3-030-24643-3_7

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-3-030-24643-3_7

Published	Publisher Name	Print ISBN
01 August 2019	Springer, Cham	978-3-030-24642-6

Online ISBN	eBook Packages
978-3-030-24643-3	Engineering Engineering (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

ScienceDirect

Biocatalysis and Agricultural Biotechnology
Volume 29, October 2020, 101761

Effect of seaweed on seed germination and biochemical constituents of *Capsicum annuum*

Mary Shamya Arokia rajan [a], Rajasekar Thriunavukkarasu [a], Jerrine Joseph [a], Wilson Aruni [a, b, c]

[a] Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, 600 119, India
[b] California University of Science and Medicine, School of Medicine, CA, USA
[c] Musculoskeletal Disease Research Laboratory US Department of Veteran Affairs, Loma Linda, CA, USA

Received 7 February 2020, Revised 19 July 2020, Accepted 17 August 2020, Available online 26 August 2020, Version of Record 6 September 2020.

Check for updates

Show less

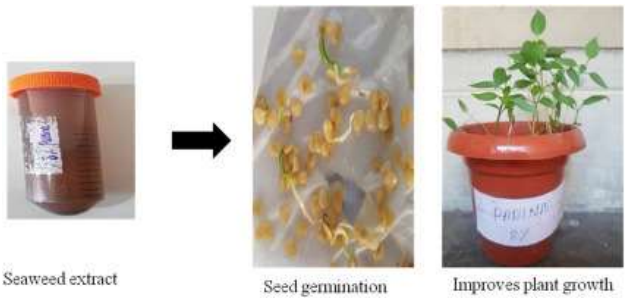
Share Cite

https://doi.org/10.1016/j.bcab.2020.101761 Get rights and content

Abstract

Macro algae or Seaweeds contain many bio active compounds which can be used in various industries including agricultural, food, and pharmaceutical. In the present study we used Padina *gymnospora, Gracilaria edulis and* Ulva *fasciata aqueous extracts used as a* seaweed biofertilizers for the seed germination of Capsicum *annuum.* The seeds were soaked in different concentrations of seaweed extract (2, 4, 6, 8 and 10%) and water (as control) for 24 h. Percentage of seed germination (GP), shoot length (SL), germination index (GI), mean germination time (MGT), seedling vigor index (SVI), germination energy (GE), pH has been studied. The minimum concentration 8% of *Padina gymnospora 6% of Gracilaria edulis* and 8% of *Ulva fasciata* enhances the seed germination rate. The above mentioned three seaweed seed germination parameters on Capsicum annuum are as follows, the shoot length was (76,70 and 85 mm in dm), germination index (0.933, 0.866 and 0.7), germination percentage (70%,65% and 60%), mean germination time (10.5,5.25 and 9), seedling vigor index (53.2%,45.55% and 51%), germination energy (70% 65% and 60%), total protein content (2%,0.1% and 10.2%), total phenol content (15%,0.09% and 12%) and antioxidant inhibition percentage (43.1%,22.4% and 24.6) in *Capsicum annuum.* Between these three seaweeds *Padina gymnospora* aqueous extract 8% reveal superior retaliation. We conclude that *Padina gymnospora* 8% may improve the germination process and phytochemical presence in *Capsicum annum.*

Graphical abstract



Download : Download high-res image (338KB) Download : Download full-size image

Introduction

In recent years the usage of chemical fertilizers has been increased to improve the production of crops. The effectiveness of this fertilizer will disturb the soil texture and fertility. And also it possesses harmful side effects to a living soul, who are working in that land and people consuming the chemical fertilizer used food items (Rengasamy et al., 2015). Overcome this motive use of biofertilizers to enhance crop growth and development in an eco-friendly manner. The cost of this chemical inorganic fertilizer will be higher than the organic fertilizer and also it will be easily available to the agronomist. Hence researchers focused to find out the finest and worthy biofertilizers to put back the chemical inorganic fertilizer.

Macro algae or seaweeds are generally present in the intertidal area of seacoast in the great amount. Based on their pigmentation production it has been divided into three main groups namely Ochrophyta, class Phaeophyceae (Brown) contain chlorophylls *a*, *c*, carotenoids and fucoxanthin, Rhodophyta (Red) pigments are chlorophyll *a*, phycobilins and carotenoids (β-carotene) and Chlorophyta (Green) contain pigments which are similar to land plants chlorophyll *a*, *b* and carotenoids. Considerably they differ not only with the color and also with the structural, biochemical attribute and cell wall composition. (Pereira, 2018). India is having 7500 km of coastline maritime rocky zone that will extensively support the growth of seaweeds. Economically seaweed placed an important role in many industries. They are the solitary source for manufacturing phytochemical including agar, carrageenan, and alginates. Seaweeds are used as living soul food, domestic feed, and fertilizers for the plant in many countries.

The extract obtained from macroalgae has numerable economic dominance in farming as enhancement of seed germination, fertilizer, and growth tonic to make headway yield and standards of veggies (Khan et al., 2009; Sarhan et al., 2011; Mattner et al., 2013; Hernández-Herrera et al., 2014). Seaweed are known for their extravagance in polysaccharides, carotenoids, dietary fiber, minerals, nutrients and other macro molecules, for example, protein, sugars, lipids, crucial unsaturated fats, indispensable and insignificant amino acids and polyphenol (Ganesan et al., 2008; Seedevi et al., 2017). Biofertilizers leads not exclusively to expanded harvest yield fully (Ansari et al., 2015), yet additionally minimize the harm to soil surface and other ecological issues that emerge from the utilization of chemical fertilizers (Buragohain et al., 2018). In this way, the utilization of biofertilizers offers financial and biological advantages, and supports the developing exertion to heighten economical agricultural frameworks (Xiang et al., 2012). Many studies showed that compounds such as minerals, nutrients, amino acids, vitamins, pigments, and polysaccharides (Calvo et al., 2014) present in the seaweed may refine seed germination, growth and development by providing them as bio-fertilizer or bio-stimulants. It enriches the plant tolerance towards extreme environment like high heat and drought condition (Zhang and Schmidt, 2000) which increases the nutrient uptake and antioxidant properties of the plant. (Turan and Köse, 2004; Verkleij, 1992). Seaweed liquid fertilizer is the supreme source of nitrogen, phosphorous, potassium, calcium and magnesium required for the normal growth of plants and also contains plant growth regulators like cytokinin, auxin, and gibberellins. (Williams et al., 1981, Nelson and Van Staden, 1984 and Rama Rao, 1991). It directly influence cellular metabolism and improves its yield. These extract has been enforced by soil drench, foliar spray, granules, seaweed paste as manure on plants, by enhancing water holding capacity and by encouraging growth of favorable soil microbes, nutrient uptake, fast seed germination, early flowering and fruit yield, abiotic stress tolerance, resist pests and diseases, bacterial and fungal pathogens (Nerissa et al., 2016). Due to the non-toxic, biodegradable, ecofriendly nature expand the usage of this fertilizer than the chemical inorganic fertilizer (Craigie, 2011).

Green chili (*Capsicum annuum*) native to southern North America and northern South America. It is generally taken as fresh in salad items in many parts of the world. They are the major ingredients of Indian, Thai, Mexican and Bangladeshi food. It contains low amount of fat and cholesterol and high amount of fiber, vitamins and minerals. Hot and spicy flavor is mainly because of the presence of capsaicin. Heat produced by the chili will help to relieve pain. They loaded with Vitamin C, Iron, Beta carotene and antioxidants, which protect the body against free radicals by acting as natural scavengers. They have zero calories but it will speed up the body metabolism within 3 h of consumption. The main goal of this paper is to compare three different seaweed extracts (*Padina gymnospora* (brown), *Gracilaria edulis* (red) and *Ulva fasciata* (green)) on the seed germination, growth, phytochemical contents and antioxidant capacity of seaweed extract treated *Capsicum annuum* by providing eco-friendly biofertilizer.

Section snippets

Chemical required

Folin- Ciocalteu reagent (ready to use), Bradford reagent (ready to use), DPPH (Diphenyl-2-picrylhydrazyl) (0.1 mM DPPH (0.039 mg of DPPH dissolved in ethanol), Gallic acid (ready to use), BSA (ready to use), mercuric chloride (0.1%) 0.1 g in 100 ml distilled water) and distilled water. Chemicals and solvents used in these experiments are of analytical grade....

Sample collection

The Seaweeds Padina gymnospora (brown), Gracilaria edulis (red) and Ulva fasciata (green) were collected from Mandapam, Rameshwaram, and...

Result

All the seaweed extracts have improved the growth of *Capsicum annuum*, in all different concentrations with minimum to the maximum level of differences. The seed germination capacity of *Capsicum annuum* is classified into three stages based on their sprouting such as slow grower, medium grower and a fast grower. All the seaweed extracts are good plant growth promoters but the concentration and the mineral content presence make them grow the plant fast, medium, and slow hence we are stating that...

Discussion

The earlier paper described that the Phaeophyta family seaweed contains more amount of macronutrients (Raguraman Vasantharaja et al., 2019). The compound level may vary in seaweed due to the temperature, season of cultivation, place of collection, salinity and the protocols used to determine the biochemical levels (Hanan and Shimaa, 2013). The PH of the extract will be altered mainly because of the presence of salt content and that will inhibit the germination process of the seed. The...

Conclusion

In the present study seaweed extract of *Padina gymnospora*, *Gracilaria edulis* and *Ulva fasciata* in different concentrations have been prepared, to study the germination process of *Capsicum annum*. *Padina gymnospora* 8% may enhance the growth and other biochemical parameters of *Capsicum annum* than control and other two seaweed extract. Products from seaweed have been increased in crop production, but the pathway used to improve the plant growth and other properties are still unknown. Hence it may...

Authors contributions

All authors have equal contribution in bringing out this research work....

CRediT authorship contribution statement

Mary Shamya Arokia rajan: Funding acquisition, Formal analysis, Data curation, Writing - original draft. **Rajasekar Thriunavukkarasu:** Conceptualization, Formal analysis, Data curation, Writing - original draft. **Jerrine Joseph:** Writing - original draft. **Wilson Aruni:** Writing - review & editing....

Declaration of competing interest

None....

Acknowledgements

All persons who have made substantial contributions to the work reported in the manuscript (e.g., technical help, writing and editing assistance, general support), but who do not meet the criteria for authorship, are named in the Acknowledgements and have given us their written permission to be named. If we have not included an Acknowledgements, then that indicates that we have not received substantial contributions from non-authors. The authors are grateful to the management of sathyabama...

References (54)

T.Y. Wang *et al.*
Bioactive flavonoids in medicinal plants: Structure, activity and biological fate
Asian J. Pharm. Sci. (2018)

R. Vasantharaja *et al.*
Influence of seaweed extracts on growth, phytochemical contents and antioxidant capacity of cowpea (*Vigna unguiculata* L. Walp)
Biocatalysis and agricultural biotechnology (2019)

P. Seedevi *et al.*
Bioactive potential and structural characterization of sulfated polysaccharide from seaweed (*Gracilaria corticata*)
Carbohydr. Polym. (2017)

W.R. Nelson *et al.*
The effect of seaweed concentrate on wheat culms
J. Plant Physiol. (1984)

P. Ganesan *et al.*
Antioxidant properties of methanol extract and its solvent fractions obtained from selected Indian red seaweeds
Bioresour. Technol. (2008)

S.B. Challen *et al.*
January. Growth of higher plants in response to feeding with seaweed extracts
Proc. Int. Seaweed Symp. (1966)

A. Arulkumar *et al.*
Phytochemical composition, in vitro antioxidant, antibacterial potential and GC-MS analysis of red seaweeds (Gracilaria corticata and Gracilaria edulis) from Palk Bay, India
Biocatalysis and agricultural biotechnology (2018)

M.F. Ansari *et al.*
Efficiency evaluation of commercial liquid biofertilizers for growth of Cicer aeritinum (chickpea) in pot and field study
Biocatalysis and Agricultural Biotechnology (2015)

AOAC (Association of Official Analytical Chemists)
Official Methods of Analysis
(1990)

AOSA (Association of Official Seed Analysts)
(1983)

View more references

Cited by (9)

Brown algae (Sargassum sp) extract prepared by indigenous microbe fermentation enhanced tomato germination parameters

2023, Biocatalysis and Agricultural Biotechnology

Show abstract

Seed germination studies on Chickpeas, Barley, Mung beans and Wheat with natural biomass and plastic waste derived C-dots

2022, Science of the Total Environment

Show abstract

Optimization of liquid fertilizer production from waste seaweed: A design of experiment based statistical approach

2022, Chemosphere

Show abstract

Noxious weed views and behaviors in Montana after 25 years of public education

2021, Invasive Plant Science and Management

AMF colonization affects allelopathic effects of Zea mays L. root exudates and community structure of rhizosphere bacteria

2022, Frontiers in Plant Science

Sustainable Approaches to Boost Yield and Chemical Constituents of Aromatic and Medicinal Plants by Application of Biostimulants

2022, Recent Advances in Food, Nutrition and Agriculture

View all citing articles on Scopus

Recommended articles (6)

Research article

In vitro germination of immature embryos for accelerating generation advancement in peppers (*Capsicum annuum* L.)

Scientia Horticulturae, Volume 170, 2014, pp. 203-210

Show abstract

Research article

Rhizobacteria improve the germination and modify the phenolic compound profile of pepper (*Capsicum annum* L.)

Rhizosphere, Volume 18, 2021, Article 100334

Show abstract

Research article

Effects of *A. nodosum* seaweed extracts on spinach growth, physiology and nutrition value under drought stress

Scientia Horticulturae, Volume 183, 2015, pp. 39-47

Show abstract

Research article

Interactive effects of plant growth-promoting rhizobacteria and a seaweed extract on the growth and physiology of *Allium cepa* L. (onion)

Journal of Plant Physiology, Volume 262, 2021, Article 153437

Show abstract

Research article

Influence of biostimulants-seed-priming on *Ceratotheca triloba* germination and seedling growth under low temperatures, low osmotic potential and salinity stress

Ecotoxicology and Environmental Safety, Volume 147, 2018, pp. 43-48

Show abstract

Research article

Preparation and evaluation of liquid fertilizer from *Turbinaria ornata* and *Ulva reticulata*

Biocatalysis and Agricultural Biotechnology, Volume 28, 2020, Article 101712

Show abstract

View full text

© 2020 Elsevier Ltd. All rights reserved.

Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

Biotechnological Communication

BBRC Bioscience Biotechnology Research Communications

Biosc.Biotech.Res.Comm. Vol 13 Number (1) Jan-March 2020 Pp-251-260

Effect of Probiotic Strain *Bacillus firmus* CAS 7 As Feed Supplement for Growth, Survival and Colour Enhancement of Smoke Angelfish *Apolemichthys xanthurus*

Rajeswari MV[1],* Annamalai N[2], Kumaran S[3], Wilson Aruni. A[4] and Balasubramanian T[5]

[1]*Centre for Aquaculture, Col. Dr. Jeppiaar Research Park, Sathyabama Institute of Science & Technology, Jeppiar Nagar, Chennai - 600119, Tamilnadu, India.*

[2]*Hawaii Natural Energy Institute, University of Hawaii at Manoa, 1680, East-West Road, Honolulu – 96822, HI, USA.*

[3]*Centre for Drug Discovery and Development, Sathyabama University, Jeppiar Nagar, Chennai - 600119, Tamilnadu, India.*

[4]*Department of Biotechnology, School of Bio and Chemical Engineering, Sathyabama University, Jeppiar Nagar, Chennai - 600119, Tamilnadu, India.*

[5]*CAS in Marine Biology, Faculty of Marine Sciences, Annamalai University, Parangipettai - 608502, Tamilnadu, India*

ABSTRACT

The effect of dietary management of probiotic micro organism, *Bacillus firmus* cas 7 on increase, survival and colour enhancement of smoke angelfish *Apolemichthys xanthurus* become evaluated. within the gift take a look at, fishes have been fed with basal diet supplemented with probiotic *B. firmus* cas 7 at 50, 100 and 150 mg kg-1 and manage basal food regimen containing no probiotic. the outcomes discovered that the basal weight-reduction plan supplemented with probiotic at one hundred fifty mg kg-1 produced considerably better weight benefit (80.95 ± 1.5g), precise increase price (0.675), survival charge (one hundred%) and feed conversion ratio (zero.ninety one) than diets supplemented with different concentrations and manage. furthermore, carotenoid content material was relatively better in fishes fed feed supplemented with probiotic in amount a hundred and fifty mg kg-1 (6.24 mg g-1) than a hundred and 50 mg kg-1 and manipulate (2.ninety one mg g-1). the experimental end result proves that the probiotic bacteria b. firmus cas 7 substantially stepped forward growth and color of smoke angelfish *Apolemicthys xanthurus* which might be useful in ornamental fish industry as feed complement. further, studies concerning the synthesis of carotenoid and appropriate dose of probiotic must be executed earlier than the use of probiotics on a massive scale to prevent any undesired consequences.

KEY WORDS: PROBIOTIC; ANGEL FISH; *BACILLUS FIRMUS*; GROWTH; CAROTENOIDS.

ARTICLE INFORMATION

*Corresponding Author: raji.shwehari@gmail.com
Received 8th Jan 2020 Accepted after revision 28th Feb 2020
Print ISSN: 0974-6455 Online ISSN: 2321-4007 CODEN: BBRCBA

Thomson Reuters ISI Web of Science Clarivate Analytics USA and Crossref Indexed Journal

NAAS Journal Score 2020 (4.31) SJIF: 2019 (4.196)
A Society of Science and Nature Publication, Bhopal India
2020. All rights reserved.
Online Contents Available at: http//www.bbrc.in/
DOI: 10.21786/bbrc/13.1/42

INTRODUCTION

Modern aquaculture demands alternatives that can sustain a healthy environment with best production practices (Ai, Xu, Mai, Xu, Wang, & Zhang, 2011). Probiotics in aquaculture have been reported to provide helpful property and the use of probiotics is an important management tool in fish culture (Balcazar, De Blas, Ruiz-Zarzuela, Cunningham, Vendrell, & Muzquiz,, 2006). The probiotics in aquaculture is use full for reduction of the use of harmful antimicrobial compounds, particularly antibiotics, and also improved the growth performance of the farmed species in an eco-friendly and sustainable manner (Wang, Xu, & Xia, 2005). Probiotics as feed supplements benefit to host with improving the value of feed, enzymatic role to digestion, inhibition of pathogenic microorganisms, medical properties, growth promoting factors and increasing immune response (Harikrishnan, Balasundaram, & Heo, 2010; Gupta, Gupta, & Dhawan, 2014; Gupta, Geetika, & Paromita, 2016). Several reports suggest that probiotics supplementation can cut down the cost of culture by improving the fish growth and feed utilization efficiency. The most typically used probiotics in aquaculture belong to gram-positive spore forming *bacillus spp.* (Wang et al., 2008; Gupta & Dhawan 2011, 2013; Gupta et al., 2014, 2016).

The genus Bacillus as putative probiotics has been used substantially as aquaculture feed components, due to its resistance to excessive temperature and excessive stress (Rengpipat, Rukpratanporn, Piyatiratitivorakul, & Menasa-veta, 2000). Nutritional supplementation of Bacillus spp. advanced the boom performance, immunity and ailment resistance of fish (Ai et al., 2011; Geng, Dong, Tan, Yang, Chi, Liu, & Liu, 2012) and giant freshwater prawn (*Macrobracium rosenbergii* (Gupta & Dhawan 2011, 2012). In rainbow trout (*Oncorhynchus mykiss*), significant development of feed conversion ratio (fcr), specific growth price (sgr) and protein performance ratio (in line with) turned into observed when the fish was fed with diets containing bacillus spp. (Merrifield et al., 2010; Gupta et al., 2014), as color of decorative species is taken into consideration as an essential factor for marketing of the product (Gouveia & Empis 2003; Gouveia & Rema 2005).

Color changes in fish are regularly related to environmental strain, and illumination might be a number one factor regulating pigment distribution via hormone regulation (Van der Salm, Martnez, Flik, & Wendelaar Bonga, 2004). Shade of fish pores and skin is predominantly depending on the presence of chromatophores containing colored pigments (Fox, 1957). The colour of fish pores and skin is generated via the absorption, mirrored image, and scattering of mild through the pigments and microstructures within the fish integument (Fujii, 2000). carotenoids are evidently going on pigments that variety in shades from yellow to crimson (Hill, 2002) which are lipid soluble pigments, are answerable for pores and skin coloration of ornamental fish, and might determine their commercial feed. The prosperity of the decorative fish industry has precipitated the indiscriminate use of antibiotics and chemo therapeutants for improvement in fitness and vitamins, and this therefore has led to the development of drug resistant traces of pathogenic microorganisms (Amabile-Cuevas, Cardenas-Garcia, & Ludgar,1995). Fishes are unable to perform de novo synthesis of carotenoids (cd) like other animals and consequently rely on nutritional supply to gain their herbal pigmentation (Goodwin, 1984).

Below in depth farming conditions and aquarium rearing, ornamental fish are fed completely with compound feeds, which ought to consequently be supplemented with carotenoids (Wang, Li, & Lin, 2008). The yellow, orange and red colors discovered in fish pores and skin are the end result of the group of carotenoid pigments, each carotenes and xanthophylls (Simpson, Katayama, & Chichester, 1981). Spore-forming bacillus species able to synthesizing carotenoid pigments and the biochemical evaluation at the carotenoids answerable for the yellow/ orange pigmentation found in bacillus sp. has been performed and the identification of the carotenoids was elucidated (Perez-Fons, Steiger, Khaneja, Bramley, Cutting, Sandmann, & Fraser, 2011).

The most typically found pigments had been yellow, orange and red. isolates have been nearly usually participants of the bacillus genus and in maximum cases have been related with recognised species consisting of *B. marisflavi*, *B. indicus, B. firmus, B. altitudinis and B. safensis*. 3 types of carotenoids were found with absorption maxima at 455, 467 and 492 nm, corresponding to the visible colors like yellow, orange and pink, respectively (Perez-Fons et al., 2011). A total of 1471 species of ornamental fish are traded globally (Wabnitz, Taylor, Green, & Razak, 2003) of which, marine angelfishes (Pomocanthidae) are the following maximum crucial organization than damsel and anemone fish (33%), consisting of approximately 25% of the overall change. *Pomacanthus* angelfish are the various most high prized of the coral reef fish (Frische, 1999; Debelius, Tanaka, & Kuiter, 2003). The marine angelfish circle of relatives pomacanthidae consists of 8 genera and 82 species global (Nelson, 2006).

Marine angelfish, *Apolemichthys xanthurus* (Bennett, 1833) are extensively dispensed at some point of the Indian ocean, in regions such as the Maldives and the east coast of India. these fish tend to stay in coral-rich areas singly or in pairs, often found outer reef facet and reef slopes at depths of 15 m with the most size approximately 25 cm, feeds mainly on sponges, sea squirts and small amounts of copepods (Rajeswari, & Balasubramanian,

2014) to the fine of our understanding, very little statistics is available on its replica and early ontogeny, possibly because of the problems of field observations because of low populace densities and massive domestic stages, and also because of the issue in breeding such species in small aquaria (Moyer, 1987). But, there was no tries had been made on increase, survival, coloration enhancement, herbal spawning and larval rearing of *Apolemichthys xanthurus* in captivity. Even though the software of probiotics has been diagnosed in aquaculture by means of some researchers global, no concerted tries have been made on these marine angel fishes to enhance boom and color till date. therefore an attempt turned into made to analyze the impact of probiotic bacteria B. firmus cas 7 on boom overall performance, survival and color enhancement of smoke angelfish *Apolemichthys xanthurus* that's one of the maximum appealing and commercially essential marine angel fish species.

MATERIAL AND METHODS

Isolation and identification of probiotic bacteria Bacillus firmus CAS7: Samples were collected from marine sediments of Parangipettai, Tamil Nadu, India and brought to the laboratory aseptically at 4 C. in nutrient broth sea water (NBSW) (salinity 34 g/l and pH 7.89), speared on nutrient marine agar 2216 (Himedia, USA) and incubated at room temperature for 24 h at 30 C. Colonies were spread on nutrient marine agar plates. After Gram staining and catalase testing, only one strain (CAS7) were showed Gram and catalase positive rods were retained. Retained strains were enriched for 24 h at 30 C in NBSW. Colonies were re-isolated on nutrient marine agar and preserved in the laboratory. The purity of such cultures was routinely checked during this work. Stock cultures were frozen at 80 C with 20% (vol/vol) glycerol.

Table 1. Morphological, physiological and biochemical characteristics of the probiotic strain CAS 7

Characteristics	Results
Shape	Rod
Gram stain	Positive
Spore formation	+
Motility	+
Glucose	+
Mannitol	+
Xylose	+
Starch Hydrolysis	+
Gelatin Hydrolysis	+
Fat Hydrolysis	+
Casein Hydrolysis	+
Catalase activity	-
Nitrate reduction	+
Indole	+
Citrate	+

Characterization of the potent strain: Pure cultured colonies were biochemically characterized (Himedia, Mumbai). The identification was confirmed by partial sequencing the 16 S rDNA by kumaran et al., 2009. The isolated bacterial DNA were extracted and 16s rDNA sequences were amplified by polymerase chain reaction (PCR) using universal primers of 8f (3'-AGAGTTTGATCCTGTGCTCAG 5') and 1490r (5'-GACTTACCAGGGTATCTAATCC-3'). The PCR products were electrophoresed on 1% agarose gel and visualized via ultraviolet transillumination. After that the PCR product was purified to remove the primer, primer-dimers and low molecular weight DNA fragments generated by nonspecific amplification. Five volume of binding buffer was mixed with one volume of PCR product and loaded into the purification column. The nucleotide sequences of the PCR products were determined by using the automated DNA sequence with forward and reverse primers (Bio serve pvt. Bio Technologies, India). The sequences was compared with all 16S rRNA gene sequence data stored NCBI by nucleotide BLAST.

Table 2. Formulation and Proximate composition of basal diet

Ingredients	g kg-1
Fish meal	600
Shrimp meal	160
Soybean meal	20
Wheat flour	140
Fish oil	40
Soybean phospholipids	20
Vitamin mineral mix	10
Proximate composition	
Dry matter	941
Crude protein	491
Crude lipids	98
Ash	96

Vitamin mineral mix (EMIX PLUS, Mumbai, India) (Quantity per kg)

Vitamin A: 22 00 000 IU; Vitamin D3: 4 40 000 IU; Vitamin B2: 800 mg; Vitamin E: 300 mg; Vitamin K: 400 mg; Vitamin B6: 400 mg; Vitamin B12: 2.4 mg; Calcium Pantothenate: 1000 mg; Nicotinamide: 4 g; Choline Chloride: 60 g; Mn: 10 800 mg; I: 400 mg; Fe: 3000 mg; Zn: 2000 mg; Cu: 800 mg; Co: 180 mg; Ca: 200 g; P: 120 g; L L-lysine: 4 g; DL-Methionine: 4 g; Selenium: 20 ppm.

Culture of probiotic bacteria B. firmus CAS 7: The strain *B. firmus* CAS 7 was cultured and prepared as described by Sun et al., (2010). 500 mL of fresh nutrient broth was seeded with 1% inoculum (1.50 x 106 CFU mL-1) and kept in a shaker incubator (200 rpm) at pH 7.5, temperature 28 oC, and salinity 30 PSU for 48 h. After incubation

period, the cells were harvested by centrifugation at 5000 xg for 10 min, washed twice with phosphate-buffered saline (pH 7.5) and re-suspended in same PBS buffer for addition to the basal diet and this was administrated for probiotic study. Cell growth was estimated by measuring optical density at 600 nm from the aliquots withdrawn at every 6 h intervals.

Preparation of control and probiotic feed: The control basal diet was formulated using the ingredients such as fish meal, shrimp meal, soya bean meal, wheat flour, fish oil and vitamin mineral mix (Table1) (Sun, Yang, Ma, & Lin, 2010).All the ingredients were dried overnight at 80° C in a hot air oven and powdered. The powdered ingredients were sieved through a fine-meshed screen (0.5 mm diameter) and mixed well. The dough was prepared by adding required amount of water with the ingredients, sterilized (autoclave at 121° C for 15 min) and incorporated with 3% (v/w) commercial vitamin mineral mix (EMIX PLUS, Mumbai, India) and pelletized using hand pelletizer to obtain 1 mm pellets. The pellets were initially sun dried and then oven dried at 60 ± 5°C for 12 hours to get moisture content. Further, they were manually broken into smaller bits and stored at room temperature in an air tightened sterile polypropylene containers.

The test feeds for the experiments were prepared by gently spraying the required amount of bacterial suspension on the control diet and mixing it part-by-part in a drum mixer to obtain a final concentration of 50, 100 and 150 mg kg-1 probiotic based on preliminary experiments in our laboratory. The probiotic cell suspensions (108cells mL-1) were added to the control diet after the dosage had been autoclaved and subsequently cooled, before pelletizing. The proximate composition (moisture, protein, ash, lipid and fibre) of all probiotic feeds and control feed were determined by the standard procedures of AOAC (1990). The viability counts (CFU/g) of the probiotic incorporate feed before and after drying was enumerated using a pour plate technique using the method described by Reddy et al., (2009) and expressed in % of survival. The probiotic strain-incorporated feeds were packed in sterile polypropylene containers and stored at 4°C for further studies.

Experimental design: The experiment was conducted at Aquaculture breeding center, Centre of Advanced Study in Marine Biology, Annamalai University, India. The juveniles of smoke angelfish, *Apolemichthys xanthurus* were obtained from MAV Breeders (Mandabam, Tamil Nadu, India) and acclimatized for 4 weeks before the trial. The feeding experiment was conducted in (20 L) rectangular fibre glass tanks, with temperature ranging from 26 - 30oC, salinity 28 -30 PSU, pH 7.4 - 7.8; and dissolved oxygen (DO) 4.2 to 5.6 mg L-1. A total of 30 fish seeds were maintained in each tank throughout the experiment (12.89 ± 0.41 g mean weight) and each treatment was conducted in triplicate. In control, the fishes were feed with prepared pellet feed alone. The probiotic bacteria of 50, 100 and 150 mg kg-1 were mixed with experimental feed. The feeding rate was about 3% of biomass per day provided in equal rations at 8.00 AM, 1.00 PM, 6.00 PM for 120 days and the excess diet was collected and dried at 60 °C, put in room temperature for 3 days to restore the natural moisture and then weighed. Feeding rate was adjusted every fortnight during sampling by batch weighing of fish in each tank after a 24 h period of starvation. At the same time, fish survival was also determined by counting the individuals in each aquarium. Experimental tanks were cleaned and water exchange was done once a week.

Growth indices: The growth parameters such as weight gain, specific growth rate (SGR), survival rate and feed conversion ratio (FCR) were assessed at 30, 60, 90 and 120 days. The weight gain (WG), specific growth rate (SGR) and feed conversion ratio (FCR) was evaluated based on standard formula as follows.

Weight gain (g) = (Final body weight (g) – Initial body weight (g)

Specific growth rate (SGR) = 100 (loge final body weight - loge initial body weight) / culture period (days)

Feed conversion ratio (FCR) = total dry feed fed (g) / total live weight gain (g)

Colour enhancement and carotenoid content estimation

The color enhancement was monitored by visual examination and estimation of carotenoid content in the skin of experimental fishes. The carotenoid content of the experimental fish skin was extracted according to the method of Torrissen & Naevdal (1984). Three fishes from each experiment were randomly sampled, anesthetised using clove oil (dissolved in 95% ethanol at 1:10) and used for carotenoid content analysis in triplicate. Briefly, 2 mg of skin were collected from both sides between the abdominal and dorsal regions of the fish and then transferred to 10 mL of pre - weighed glass tubes after the fat layer had been removed from the skin and ground well with acetone containing anhydrous sodium sulphate and made up to 10 mL with acetone. The samples were stored for 3 days at 4oC in a refrigerator, and then extracted three times till no further colour could be obtained and centrifuged at 5000 xg for 5 min. The total carotenoid content of the samples was determined using spectrophotometer (Shimadzu, UV mini 1240) using extinction coefficients (E1%, 1 cm) of 2000 for astaxanthin (Hata, & Hata, 1971) at 475 nm, and 2500 for carotenoids from alfalfa at 450 nm (Schiedt, & Jensen, 1995).

Data analysis: In the present study, all the data were analyzed by statistical methods. The two way analysis of variance (ANOVA) was performed using SPSS (Statistic Package for social science) version 11.5 software to determine the significant differences among means. For all tests, a criterion of P<0.01 was used to determine statistical significance.

RESULTS AND DISCUSSION

Based on the morphological, physiological, and biochemical characteristics, the strain CAS 7 is a Gram-positive and endospore-forming bacillus, with catalase but without oxidase, which grows in both aerobic and anaerobic environments (Table 1). Further, 16S rRNA gene sequencing and BLAST analysis confirmed that the strain belongs to *B. firmus* and designated as *B. firmus* CAS 7 (GenBank accession no. HQ116805).The 16S rRNA gene sequences of the probiotic strain CAS 7 obtained from the present study was deposited in NCBI with accession number HQ116811.

The results revealed that the weight of fishes increased with the increase in days of culture in all the experimental groups (Table.2). At the end of the experiment, it was found that there was significant weight gain in all the three experiments where fishes were fed with probiotic-supplemented basal feed. The weight gain was comparatively less in the control group. The weight gain was the highest in group 150 mg kg-1(80.95 g) followed by group 100 (67.26 g), 50 (57.37 g) and control groups (45.81 g). The specific growth rate (SGR) was significantly (P<0.01) higher in 150 mg kg-1 group (0.675) compared with fishes in the group of 100 mg kg-1 (0.527), 50 mg kg-1 (0.457) and control (0.382). Furthermore, the most advantageous value for conversion ratio (FCR) classified as follows: 150 > 100 > 50 > control group (Table 3). The survival rate of the fishes was 80 - 100% (in control and 150 mg kg-1group, respectively) (P<0.01).

The results of the carotenoid analysis revealed that the initial carotenoid content of the fish skin varied between 1.24 and 1.28 mg g-1 in all experimental groups and control and increased gradually with increasing days of culture. At the end of the experimental period, carotenoid content was 2.91 - 6.24 mg g-1 (in control and 150mg kg-1 group) (Table 3) (P<0.01) (Table.3). Thus, these results depicted that the carotenoid content of fish group fed with feed was supplemented with 150 mg kg-1 of probiotic was the highest. The research on the use of probiotics in aquatic animals has received heightened attention with the demand for environment friendly health management in aquaculture (Geng et al., 2011). A growing number of studies have demonstrated the ability of probiotics to increase the growth rate and / or feed utilization of farmed aquatic animals (Carnevali et al., 2004; Wang et al., 2008). The benefits of probiotics in fish farming are improvements of growth performances, immunity and pathogen exclusions (Qi, et al., 2009; Sun et al., 2010). Most of the probiotic studies have focused on use of *Lactobacillus* and *Bacillus spp.* (Lee, 2013).

Bacillus species are gaining more and more importance and are widely used in aquaculture due to their longer stability, easy preparedness, antagonistic effects on pathogens and enhancement of immunity (Hong et al., 2005; Gupta et al., 2014). Earlier studies suggested that the Bacillus species significantly improved the growth in Catla catla (2 × 105 B. circulans PB7 cells per 100 g feed) (Bandyopadhyay & Mohapatra ,2009), Labeo rohita (1.5 x 105 *B. circulans*) (Ghosh et al., 2003), Macrobrachium rosenbergii (*Artemia salina nauplii* with *B. subtilis* (108cells mL-1) (Keysami et al., , 2007) and Penaeus monodon (Rahiman, et al.2010). Although several studies have been undertaken to study the beneficial effects of probiotic bacteria on growth and survival rate of various organisms, no studies have been attempted on marine ornamental fishes, especially on marine angel fish, *Apolemichthys xanthurus*. Hence, an attempt was made in this study to ascertain the effect of probiotic, Bacillus firmus CAS 7 as feed supplement on growth, survival and colour enhancement in marine angelfish *Apolemichthys xanthurus*.

In the present study, the fishes were fed basal diet supplemented with probiotic *B. firmus* CAS 7 at 50, 100 and 150 mg kg-1 and the growth performance such as weight gain, survival rate and color enhancement was evaluated for 120 days. Results suggested that the growth, weight gain and SGR were comparatively higher in fish group fed with feed supplemented 150 mg kg-1 of probiotic than other experimental groups and control. Other reports also reported that the probiotic (Bacillus *sp.*) supplemented diet (bio encapsulated *Artemia* nauplii by 3×108 CFU/ L) significantly increased the weight, length and SGR of fish when compared to the control diet. Several studies suggested that the probiotic supplementation has significantly increased the weight gain and SGR in *Rachycentron canadum* (Geng et al., 2012), Labeo rohita (Giri, Sukumaran, & Oviya, 2013) (Giri et al., 2013), *Oreochromis niloticus* (Aly et al., 2008), *Epinephelus coioides* (Sun et al., 2010) and *Larimichthys crocea* (Ai et al., 2011).

The results suggested that the FCR was lower in the fish group fed with feed supplemented 150 mg kg-1 probiotic followed by 100 and 50 150 mg kg-1 and control. It seems that the reduction in FCR of fishes in experimental groups revealed dietary nutrients were utilized more efficiently when the diet was supplemented with probiotics. Similarly, Neja, Rezaei, Takami, Lovett, Mirvaghefi, & Shakouri (2006) reported that *Bacillus spp.* could be used to increase the digestive enzyme activity, survival and growth in the Indian white shrimp.Fishes

are incapable of biosynthesizing carotenoids, so diet is their sole source as only plants, bacteria, fungi and algae have the capacity for its synthesis (Geng et al., 2011). Many reports have demonstrated that skin color change over time depended on the level of carotenoid in the diet and differed among species (Chatzifotis, Pavlidis, Jimeno, Vardanis, Sterioti, & Divanach, 2005; Dharmaraj, & Dhevendaran, 2011; Ho, Zong, & Lin, 2014). The results of the study on color enhancement suggested that the carotenoid content in the skin of fishes fed with probiotic maintained coloration during periods of social interaction, suggesting that the probiotic Bacillus firmus CAS7 may play important roles in maintaining fish skin coloration. Steiger, (2012)& Liu et al., (2009) were suggested that carotenoids are commonly found in pigmented bacteria which are known to have a positive role in the intermediary metabolism of fish that could enhance nutrient utilization.The research on the use of probiotics in aquatic animals has received heightened attention with the demand for environment friendly health management in aquaculture (Geng et al., 2011).

A growing number of studies have demonstrated the ability of probiotics to increase the growth rate and / or feed utilization of farmed aquatic animals (Carnevali et al., 2004; Wang et al., 2008). The benefits of probiotics in fish farming are improvements of growth performances, immunity and pathogen exclusions (Qi et al., 2009; Sun et al. 2010). Most of the probiotic studies have focused on use of Lactobacillus and Bacillus spp. (Cabello et al., 2013; Lee, Kim, Song, Oh, Cha, Jeong, Heo, Kim, & Lee, 2013). Bacillus species are gaining more and more importance and are widely used in aquaculture due to their longer stability, easy preparedness, antagonistic effects on pathogens and enhancement of immunity (Hong, Le Duc, Cutting, 2005; Gupta et al., 2014).

Table 3. Growth performance and survival rate of smoke angel fish Apolemichthys xanthurus fed with basal diet (control -without probiotic) and basal diet supplemented with 50, 100 and 150 mg kg-1 of probiotic strain B. firmus CAS7. Values are presented in mean ± SD, n = 3, FCR - feed conversion ratio; SGR - specific growth rate (% d-1)

Days of culture	Growth Parameters	Control	50 mg kg^{-1}	100 mg kg^{-1}	150 mg kg^{-1}
0 - 30	Initial weight (g)	12.33±0.38	12.57±0.29	12.63±0.32	12.89±0.41
	Final weight(g)	20.27±1.60	25.94±1.20	29.93±1.40	30.13±1.10
	Weight gain (g)	7.94±1.60	13.37±0.29	17.3±0.30	17.24±1.20
	SGR	0.265±0.03	0.446±0.04	0.577±0.02	0.641±0.03
	FCR	2.33±0.04	1.38±0.03	1.07±0.03	1.07±0.02
	Survival rate (%)	80	90	90	100
	Carotenoid content (mg g-1)	1.24±0.06	1.28±0.08	1.25±0.04	1.25±0.06
0 - 60	Initial weight (g)	12.33±0.38	12.57±0.29	12.63±0.32	12.89±0.41
	Final weight(g)	32.83±1.50	39.92±1.60	45.57±1.10	51.56±1.30
	Weight gain (g)	20.5±1.50	27.35±1.20	32.94±1.40	38.67±1.40
	SGR	0.342±0.03	0.456±0.02	0.549±0.01	0.763±0.02
	FCR	1.80±0.05	1.35±0.02	1.12±0.03	0.96±0.05
	Survival rate (%)	80	90	100	100
	Carotenoid content (mg g-1)	2.54±0.03	3.14±0.03	3.79±0.05	4.38±0.06
0 - 90	Initial weight (g)	12.33 ± 0.38	12.57 ± 0.29	12.63 ± 0.32	12.89 ± 0.41
	Final weight(g)	46.24±0.80	55.32 ±1.50	61.02 ±1.30	72.56 ± 0.90
	Weight gain (g)	33.91±1.60	42.75 ±1.50	48.39±1.30	59.67±0.90
	SGR(% day-1)	0.377±0.04	0.475±0.02	0.552±0.05	0.730±0.02
	FCR	1.64±0.05	1.30±0.03	1.15±0.06	0.93±0.01
	Survival rate (%)	80	90	90	100
	Carotenoid content (mg g-1)	2.76±0.05	3.55±0.07	4.01±0.04	5.69±0.004
0 - 120	Initial weight (g)	12.33±0.38	12.57±0.29	12.63±0.32	12.89±0.41
	Final weight(g)	58.14±1.3	69.94±1.20	79.89±0.90	93.84±1.40
	Weight gain (g)	45.81±1.4	57.37±1.30	67.26±1.40	80.95±1.50
	SGR	0.382±0.03	0.457±0.03	0.527±0.02	0.675±0.04
	FCR	1.61±0.03	1.29±0.03	1.10±0.04	0.91±0.02
	Survival rate (%)	80	90	90	100
	Carotenoid content (mg g-1)	2.91±0.08	4.26±0.06	5.74±0.03	6.24±0.05

Earlier studies suggested that the Bacillus species significantly improved the growth in Catla catla (2 × 105 B. Circulans PB7 cells per 100 g feed) (Bandyopadhyay & Mohapatra ,2009), Labeo rohita (1.5 x 105 *B. circulans*) (Ghosh et al., 2003), Macrobrachium rosenbergii (*Artemia salina nauplii* with *B. subtilis* (108cells mL-1) (Keysami et al., 2007) and *Penaeus monodon* (Rahiman et al., 2010). Although several studies have been undertaken to study the beneficial effects of probiotic bacteria on growth and survival rate of various organisms, no studies have been attempted on marine ornamental fishes, especially on marine angel fish, *Apolemichthys xanthurus*. Hence, an attempt was made in this study to ascertain the effect of probiotic, *Bacillus firmus* CAS 7 as feed supplement on growth, survival and colour enhancement in marine angelfish *Apolemichthys xanthurus*.

In the present study a *B. firmus* CAS 7 isolated from marine environment was cultured in nutrient broth for 48 h and used as feed supplement to evaluate the improvement in growth and colour enhancement on marine smoke angel fish. The growth kinetic study revealed that the maximum cell growth of the probiotic bacterium was achieved at logarithmic phase (24th h). Elayaraja et al., (2011) studied the effect of amylase produced by *B. cereus* on the growth of polycheate and maximum cell growth as well as enzyme production was found at late logarithmic phase (36th h). The short incubation period for achieving maximum growth makes this a potential probiotic candidate species which could be used in the aquaculture industry. The dietary supplementation of *B. firmus* CAS 7 in the present study exhibited good growth performance and significantly increased survival rate of *Apolemichthys xanthurus*. In a previous study, Bacillus administration has also been shown to increase survival by enhancing resistance to pathogens by acting both cellular and humoral immune defence in shrimp and prawn (Rengpipat et al., 2000; Gupta, & Dhawan, 2013).

Abraham et al., (2007) suggested that the probiotics greatly helps in achieving natural resistance and high survivability of fish. In the present study, the fishes were fed basal diet supplemented with probiotic *B. firmus* CAS 7 at 50, 100 and 150 mg kg-1 and the growth performance such as weight gain, survival rate and color enhancement was evaluated for 120 days. Results suggested that the growth, weight gain and SGR were comparatively higher in fish group fed with feed supplemented 150 mg kg-1 of probiotic than other experimental groups and control. Jafaryan et al., (2008) also reported that the probiotic (*Bacillus sp*.) supplemented diet (bio encapsulated *Artemia nauplii* by 3×108 CFU/ L) significantly increased the weight, length and SGR of fish when compared to the control diet. Several studies suggested that the probiotic supplementation has significantly increased the weight gain and SGR in *Rachycentron canadum* (Geng et al., 2012), Labeo rohita (Giri et al., 2013), *Oreochromis niloticus* (Aly et al., 2008), *Epinephelus coioides* (Sun et al., 2010) and *Larimichthys crocea* (Ai et al., 2011).

The results suggested that the FCR was lower in the fish group fed with feed supplemented 150 mg kg-1 probiotic followed by 100 and 50 150 mg kg-1 and control. It seems that the reduction in FCR of fishes in experimental groups revealed dietary nutrients were utilized more efficiently when the diet was supplemented with probiotics. Similarly, Nejad et al., (2006) reported that Bacillus spp. could be used to increase the digestive enzyme activity, survival and growth in the Indian white shrimp. Fishes are incapable of biosynthesizing carotenoids, so diet is their sole source as only plants, bacteria, fungi and algae have the capacity for its synthesis (Geng et al., 2011). Many reports have demonstrated that skin color change over time depended on the level of carotenoid in the diet and differed among species (Chatzifotis et al., 2005; Dharmaraj, & Dhevendaran, 2011; Ho et al., 2014). The results of the study on color enhancement suggested that the carotenoid content in the skin of fishes fed with probiotic maintained coloration during periods of social interaction, suggesting that the probiotic Bacillus firmus CAS7 may play important roles in maintaining fish skin coloration. Steiger, (2012) & Amar et al., (2001) were suggested that carotenoids are commonly found in pigmented bacteria which are known to have a positive role in the intermediary metabolism of fish that could enhance nutrient utilization and may ultimately result in improved growth. in improved growth.

CONCLUSION

The results confirmed that Apolemicthys xanthurus fed with 150 mg kg-1 probiotic bacteria B. firmus CAS 7 exhibited significantly improved growth performance, survival and colour enhancement. Therefore, 150 mg kg-1 of probiotic B. firmus CAS 7 should be used for formulating nutritionally balanced diet of marine angel fish for its better growth and survival. The mode and mechanism of actions about carotenoid content increase in the fish skin and carotenoid production by this probiotic strain need further investigation.

ACKNOWLEDGEMENTS

The authors are greatly thanked to the authorities of Sathyabama Institute of Science and Technology and Annamalai University for providing the facilities. The author, M.V. Rajeswari is greatly acknowledged to University Grants Commission (UGC), Government of India for financial support through UGC-Postdoctoral fellowship for women.

REFERENCES

Abraham, T.J., Babu, C.H.S., Mondal, S., & Banerjee, T. (2007). Effects of dietary supplementation of commercial human probiotic and antibiotic on the growth rate and content of intestinal microflora in ornamental fishes. Bangladesh Journal of Fisheries Research, 11, 57–63. http://dx.doi.aquaticcommons.org/id/eprint/18902

Ai, Q., Xu, H., Mai, K., Xu, W., Wang, J., & Zhang, W. (2011). Effects of dietary supplementation of Bacillus subtilis and fructooligo saccharide on growth performance, survival, non-specific immune response and disease resistance of juvenile large yellow croaker, Larimichthys crocea. Aquaculture, 317, 155-161. http://dx.doi.org/10.1016/j.aquaculture.2011.04.036

Aly, S.M., Mohamed, M.F., & John, G. (2008). Effect of probiotics on the survival, growth and challenge infection in Tilapia nilotica (Oreochromis niloticus). Aquaculture Research, 39, 647-656. https://dx.doi.org/10.1111/j.1365-2109.2008.01932.x

Amabile-Cuevas, C.F., Cardenas-Garcia, M., & Ludgar, M. (1995). Antibiotic resistance. Animal Science, 83, 320-329. https://dx.doi.org/10.1111/j.1365-2109.2008.01932.x

Amar, E.C., Kiron, V., Satoh, S., & Watanabe, T. (2001). Influence of various dietary synthetic carotenoids on bio-defence mechanisms in rainbow trout, Oncorhynchus mykiss (Walbaum). Aquaculture Research, 32, 162-163. https://dx.doi.org/10.1046/j.1355-557x.2001.00051.x

AOAC (1990). Official Methods of Analysis of the Association of Official Analytical Chemists, Vol. 1, 14th edn. 1102 pp.

Balcazar, J.L., De Blas, I., Ruiz-Zarzuela, I., Cunningham, D., Vendrell, D., & Muzquiz, J.L. (2006). The role of probiotics in aquaculture. Veterinary Microbiology, 114, 173-186. https://dx.doi:10.1016/j.vetmic.2006.01.009

Bandyopadhyay, P., & Mohapatra, P.K.D. (2009). Effect of a probiotic bacterium Bacillus circulans PB7 in the formulated diets: on growth, nutritional quality and immunity of Catla catla (Ham.). Fish Physiology and Biochemistry, 35, 467-478. https://dx.doi 10.1007/s10695-008-9272-8

Carnevali, O., Zamponi, M.C., Sulpizio, R., Rollo, A., Nardi, M., Orpianesi, C., Silvi, S., Caggiano, M., Polzonetti, A. M., & Cresci, A. (2004). Administration of probiotic strain to improve seabream wellness during development. Aquaculture International, 12, 377-386. https://dx.doi:10.1023/b:aqui.0000042141.85977.bb

Chatzifotis, S., Pavlidis, M., Jimeno, C.D., Vardanis, G., Sterioti, A., & Divanach, P. (2005). The effect of different carotenoid sources on skin coloration of cultured red porgy (Pagrus pagrus). Aquaculture Research, 36, 1517 – 1525. https://dx.doi:10.1111/j.1365-2109.2005.01374.x

Debelius, H., Tanaka, H., Kuiter, & R. H. (2003). Angelfishes: A Comprehensive Guide to Pomacanthidae. TMC Publishing, Chorleywood, UK, 208 pp.

Dharmaraj, S., & Dhevendaran, K. (2011). Application of microbial carotenoids as a source of colouration and growth of ornamental fish Xiphophorus helleri. World Journal of Fish and Marine Sciences, 3(2),137-144. Retrieved form https://idosi.org/wjfms/wjfms3(2)11/10.pdf

Duc, L. H., Fraser, P., & Cutting, S.M. (2006). Carotenoids present in halotolerant Bacillus spore formers. FEMS Microbiology Letter, 255, 215-224. https://dx.doi:10.1111/j.1574-6968.2005.00091.x

Ege, (1939). A revision of the genus Anguilla Shaw, a systematic, phylogenetic and geographical study, Dana Report, 16, 256 pp

Elayaraja, S., Annamalai, N., Murugesan, P., Mayavu, P., & Balasubramanian, T. (2011). Effect of amylase on growth, survival and proximate composition of the Polychaete, Perinereis cultrifera (Grube, 1840). Aquaculture Nutrition, 17, 627-633. http://dx.doi.10.1111/j.1365-2095.2011.00864.x

Fox, D.L. (1957). The pigments of fishes. Academic Press Inc, New York, USA.

Frische, J. (1999). Marine Ang7elfishes. TFH Publications, Neptune, NJ, USA, 64 pp.

Fujii, R. (2000). The regulation of motile activity in fish chromatophores. Pigment Cell Research, 13, 300-319. http://dx.doi.10.1034/j.1600-0749.2000.130502.x

Geng, X., Dong, B., Tan, Q., Yang, S., Chi, H., Liu. (2011). Effects of dietary chitosan and Bacillus subtilis on the growth performance, non-specific immunity and disease resistance of cobia, Rachycentron canadum. Fish Shellfish Immunology, 31, 400 – 406. http://dx.doi.10.1016/j.fsi.2011.06.006

Geng, X., Dong, X, H., Tan, B.P., Yang, Q. H., Chi, S.Y., Liu, H.Y., & Liu, X.Q. (2012). Effects of dietary probiotic on the growth performance, non-specific immunity and disease resistance of cobia, Rachycentron canadum. Aquaculture Nutrition, 18, 46-55. https://dx.doi.org/10.1111/j.1365-2095.2011.00875.x

Ghosh, K., Sen, S.K., Ray, A.K. (2003). Supplementation of an isolated fish gut bacterium, Bacillus circulans, in formulated diets for rohu, Labeo rohita, fingerlings. Israeli J. Aquaculture Nutrition, 55, 13-21. http://dx.doi.10.1111/j.1365-2095.2011.00875.x

Giri, S.S., Sukumaran, V., & Oviya, M. (2013). Potential probiotic Lactobacillus plantarum VSG3 improves the growth, immunity, and disease resistance of tropical freshwater fish Labeo rohita. Fish Shellfish Immunology, 34, 660-666. http://dx.doi.org/10.1016/j.fsi.2012.12.008

Gouveia, L., & Empis, J. (2003). Relative stabilities of microalgal carotenoids in microalgal extracts, biomass and fish feed: effect of Storage Condition. Innovative Food Science Emerging Technology, 4, 227-233. http://dx.doi.10.1016/S1466-8564(03)00002-X

Gouveia, L., & Rema, P. (2005). Effect of microalgal biomass concentration and temperature on ornamental goldfish (Carassius auratus). Aquaculture Nutrition, 11, 19 – 23. http://dx.doi.10.1111/j.1365-2095.2004.00319.x

.Gupta, A., & Dhawan, A. (2011). A Effect of

supplementing probiotics (Prosol) on the performanceof giant freshwater prawn (Macrobrachium rosenbergii) juveniles. Indian Journal of Animal Nutrition, 28, 457-63. http://dx.doi:10.1111/j.1365-2109.2007.01787.x

Gupta, A., Dhawan, A. (2012). Effect of dietary probiotic Improval (Lactobacillus sporogenes and Saccharomyces cerevisiae) on growth and feed utilization of Macrobrachium rosenbergii post larvae. Animal Nutrition Feed Techechnology, 12, 209-17. Retrived by https://www.indianjournals.com/ijor.aspx?target=ijor:anft&t volume=12&t issue=2&article=007

Gupta, A., Dhawan, A. (2013). Probiotic based diets for freshwater prawn Macrobrachium rosenbergii (de Man). Indian Journal of Fisheries, 60, 103–109. Retrieved from http://epubs.icar.org.in/ejournal/index.php/IJF/article/view/ 17914/12929

Gupta, A., Gupta, P., Dhawan, A. (2014). Dietary supplementation of probiotics affects growth, immune response and disease resistance of Cyprinus carpio fry. Fish and Shellfish Immunology, 41, 113–119. http://dx.doi.10.1016/j.fsi.2014.08.023.

Gupta, A., Geetika, V., Paromita, G. (2016). Growth performance, feed utilization, digestive enzyme activity, innate immunity and protection against Vibrio harveyi of freshwater prawn,Macrobrachium rosenbergii fed diets supplemented with Bacillus coagulans. Aquaculture international, http://dx.doi.10.1007/s10499-016-9996-x.

Harikrishnan, R., Balasundaram, C., & Heo, M.S. (2010). Lactobacillus sakei BK19 enriched diet enhances the immunity status and disease resistance to streptococcosis infection in kelp grouper. Epinephelus bruneus. Fish and Shellfish Immunology, 29, 1037-1043. http://dx.doi.10.1016/j.fsi.2010.08.017

Hata, M., & Hata, M. (1971). Carotenoid pigments in goldfish (Carassius auratus) II. Colour change and carotenoid pigment composition. International Journal of Biochemistry, 2, 182-184. http://dx.doi.10.1016/0020-711X(71)90210-2

Hill, G.E. (2002). A Red Bird in a Brown Bag: The Function and Evolution of Colorful Plumage in the House Finch. Oxford University Press, New York. 35 pp

Ho, A.L.F. C., Zong, S., & Lin, J. (2014). Skin color retention after dietary carotenoid deprivation and dominance mediated skin coloration in clown anemonefish, Amphiprion ocellaris. AACL Bioflux 7(2), 103-115. Retrieved by https://www.bioflux.com.ro/docs/2014.103-115.pdf

Hong, H.A., Le Duc, H., Cutting, S.M. (2005). The use of bacterial spore formers as probiotics. FEMS Microbiol Reviews, 29, 813–835. http://dx.doi.10.1016/j.femsre.2004.12.001

Keysami, M.A., Saad, C.R., Sijam, K., Daud, H.M., & Alimon, A. R. (2007). Effect of Bacillus subtilis on growth development and survival of larvae Macrobrachium rosenbergii (de Man) Aquaculture Nutrition, 13, 131-136. http://dx.doi- 10.1111/j.1365-2095.2007.00463.x

Lee, B.J., Kim, S.S., Song, J.W., Oh, D.H., Cha, J.H., Jeong, J.B., Heo, M.S., Kim, K, W., & Lee, K. J. (2013). Effects of dietary supplementation of citrus by-products fermented with a probiotic microbe on growth performance, innate immunity and disease resistance against Edwardsiella tarda in juvenile olive flounder, Paralichthys olivaceus (Temminck and Schlegel). Journal of Fish Diseases, 36, 617–628. http://dx.doi. 10.1111/jfd.12035

Matsuno., T. (2001). Aquatic animal carotenoids. Fisheries Science, 67, 771–783.

Moyle, P.B., Cech, J.J. (2004). Fishes, an introduction to ichthyology. , Upper Saddle River, New Jersey, USA. Prentice-Hall, 590 pp.

Merrifield, D. L., Dimitroglou, A., Bradley, G., Baker, R.T.M. & Davies, S. J. (2010). Probiotic applications for rainbow trout (Oncorhynchus mykiss Walbaum) I. Effects on growth performance, feed utilization, intestinal microbiota and related health criteria. Aquaculture Nutrition, 16, 496-503. http://dx.doi.10.1111/j.1365-2095.2009.00689.x

Moyer, J.T. (1987). Social organization and protogynous hermaphroditism in marine angelfishes (Pomacanthidae). In: Sex Change in Fishes (ed. byA. Nakazono and T. Kuwamura), Tokai University Press, 120-147 pp.

Nejad, S., Rezaei, M. H., Takami, G.A., Lovett, D.L., Mirvaghefi, A.R., & Shakouri M. (2006). The effect of Bacillus spp. bacteria used as probiotics on digestive enzyme activity, survival and growth in the Indian white shrimp Fenneropenaeus indicus. Aquaculture, 252, 516-524. http://dx.doi.10.1016/j.aquaculture.2005.07.021

Perez-Fons, L., Steiger, S., Khaneja, R., Bramley, P.M., Cutting, S.M., Sandmann, G., & Fraser, P.D. (2011). Identification and the developmental formation of carotenoid pigments in the yellow/orange Bacillus spore-formers. Journal of Biochimica et Biophysica Act, 1811, 177-185. http://dx.doi.10.1016/j.bbalip.2010.12.009

Qi, Z.Z., Zhang, X.H., Boon, N., & Bossier, P. (2009). Probiotics in aquaculture of China-current state, problems and prospect. Aquaculture, 290, 15-21. http://dx.doi.10.1016/j.aquaculture.2009.02.012.

Rahiman, K. M. M., Jesmi, Y., Thomas, A.P., & Hatha, A.A.M. (2010). Probiotic effect of Bacillus NL110 and Vibrio NE17 on the survival, growth performance and immune response of Macrobrachium rosenbergii (de Man). Aquaculture Research, 41, 120-134. http://dx.doi.10.1111/j.1365-2109.2009.02473.x.

Rajeswari, M.V., & Balasubramanian, T. (2014). Distribution, diversity and taxonomy of marine angelfishes (Pomacanthidae) of Tamil Nadu, Southeast coast of India. International Journal of Fisheries and Aquaculture, 6, 20-31. http://dx.doi. 10.5897/ijfa12.069

Rengpipat, S., Rukpratanporn, S., Piyatiratitivorakul, S., & Menasa-veta., P. (2000). Immunity enhancement on black tiger shrimp (Penaeus monodon) by a probiont bacterium (Bacillus S11). Aquaculture, 191, 271-

288. Retrieved from http://www.elsiever.ni/locate/aquaonline/2000

Schiedt, K., Jensen, S.L. (1995). Isolation and analysis. In: Britton, G., Liaaen-Jensen, S., Pfander, E. (Eds.), Carotenoids: Isolation and Analyses, vol. 1A. Birkha¨user, Basel, Switzerland, 108 pp

Simpson, K.L., Katayama, T., & Chichester, C.O. (1981). Carotenoids in fish feeds. In: Carotenoids as Colorants and Vitamin A Precursors (Bauernfeind, J.C. ed.) volume 4: Academic Press, New York. 463 pp

Steiger, S., Perez-Fons, L., Fraser, P.D. & Sandmann, G. (2012). Biosynthesis of a novel C30 carotenoid in Bacillus firmus isolates, Journal of Applied Microbiology, 113, 888-895. http://dx.doi.10.1111/j.1365-2672.2012.05377.x

Storebakken, T. P., Foss, K., Schiedt, Austreng, E., & Jensen , S.L. (1987). Carotenoids in the diets for salmonids IV. Pigmentation of Atlantic salmon with astaxanthin, astaxanthin dipalmitate and canthaxanthin. Aquaculture,65(3-5), 279-292. https://dx.doi.org/10.1016/0044-8486(87)90241-9

Sun, Y.Z., Yang, H.L., Ma, R.L, & Lin, W.Y. (2010). Probiotic applications of two dominant gut Bacillus strains with antagonistic activity improved the growth performance and immune responses of grouper Epinephelus coioides. Fish Shellfish Immunology, 29, 803-809. http://dx.doi.10.1016/j.fsi.2010.07.018

Van der Salm, A.L., Martnez, M., Flik, G., & Wendelaar Bonga, S.E. (2004). Effects of husbandry conditions on the skin colour and stress response of red porgy, Pagrus pagrus. Aquaculture, 241: 371-386.http://doi.10.1016/j.aquaculture.2004.08.038.

Wabnitz, C., Taylor, M., Green, E., Razak, T. (2003). From ocean to aquarium. UNEP-WCMC, Cambribge, U.K. 64 pp

Wang, Y.B., Xu, Z.R., & Xia, M.S. (2005). The effectiveness of commercial probiotics in Northern White Shrimp (Penaeus vannamei L.) ponds. Fisheries Science, 71, 1034-1039. http://dx.doi.10.1111/j.1444-2906.2005.01061.x

Wang, Y.B., Li, J.R., & Lin, J. (2008). Probiotics in aquaculture: challenges and outlook. Aquaculture, 281, 1-4. http://dx.doi.10.1016/j.aquaculture.2008.06.002

Highlights

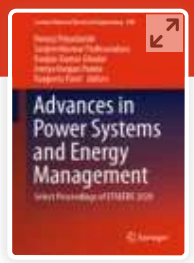
The fishes were fed basal diet supplemented with probiotic of B. firmus CAS 7 at 50, 100 and 150 mg kg-1

The growth performance such as weight gain, survival rate and color enhancement was evaluated for 120 days.

The probiotic effect of fish growth and disease resistant were showed promising activity.

Springer Link

Search Log in



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Power Systems and Energy Management** pp 575–584

EDA-PEGASIS: A Balanced Energy Aware Routing Approach for Sensor Network to Reduce Cognitive Networking Complexities in Wireless Medium

S. Sreeram & G. Nagarajan

Conference paper | First Online: 21 January 2021

300 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 690)

Abstract

The area of Wireless Sensor Network (WSN) unwraps lots of challenges to the field of cognitive networking largely, in the modern era of communication and network engineering. The basis step to tackle the issue related to listening/receiving the data is to introduce effective methodologies in the way of communicating with the nodes in the network. The paper proposes an improvement to the design of WSN formation,

where the battery operated Wireless sensor nodes (or nodes)transmits data continously to the Base station (or sink)addressing the challenges in improving QoS parameters and Energy management of nodes. In this paper, we propose a modified version of PEGASIS (Power-Efficient Gathering in Sensor Information System) protocol namely EDA-PEGASIS (Euclidean Distance Assisted – PEGASIS) to improve the lifetime of network by creating partial chains in data gathering after initial run and re-electing the leader node (LN) after energy level of the transmitting node falls below a threshold value by including Euclidean distance as a constraint. The nodes once acted as a LN will not be re-elected until energy level of all the nodes nearing to sink are below the threshold value. As a result, the average energy spent by each node per round is reduced which in turn lower the bandwidth requirement thereby achieving less power dissipation benchmark to increase efficiency and lifetime of the network. Further, this design of communication of nodes using EDA-PEGASIS will help to reduce the computational complexities in Cognitive Network.

Keywords

EDA-PEGASIS **Cognitive network**

Wireless sensor networks

Energy aware routing **Bandwidth**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7504-4_57
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 245.03
Price includes VAT (India)

- ISBN: 978-981-15-7504-4
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 299.99
Price excludes VAT (India)

- ISBN: 978-981-15-7743-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 299.99
Price excludes VAT (India)

- ISBN: 978-981-15-7503-7
- Dispatched in 3 to 5 business days

- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. S.K. Singh, M. Singh, D. Singh, A survey of energy-efficient hierarchical cluster-based routing in wireless sensor networks. Int. J. Adv. Netw. Appl. (IJANA) 2(02), 570–580 (2010)

2. J. Rana, S Vhatkar, M. Atique, Comparative study of PEGASIS and PDCH protocols in wireless sensor network, Int. J. Comput. Appl. 0975–8887 (2015)

3. N. Jamal, A.E. Al-karaki, Kamal, Routing techniques in wireless sensor networks: a survey. IEEE Wireless Commun **11**:6{28} (2004)

4. A.H. Mohajerzadeh, M.H. Yaghmaee, An efficient energy aware routing protocol for real time traffic in wireless sensor networks, in *ICUMT'09. International Conference on. IEEE Ultra-Modern Telecommunications and Workshops, 2009* (2009)

5. S. Lindsey, C.S. Raghavendra, PEGASIS: power-efficient gathering in sensor information

systems, in *IEEE proceeding on aerospace* vol. 3, pp. 1125–1130 (2002)

6. A. Devasena, B. Sowmya, A study of power and energy efficient clustering protocols in wireless sensor networks. Int. J. Adv. Res. Comput. Sci. Manage. Stud. **1**(6) (2013)

7. Y. Liu, H. Ji, G. Yue, An energy- efficient pegasis-based enhanced algorithm in wireless sensor networks. Technology Forum, China Communications, August 2006

8. F. Sen, B. Qi, L. Tang, An improved energy-efficient PEGASIS-based protocol in wireless sensor networks, in Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), vol. 4 (Shanghai, 2011), pp. 2230–2233

9. S.P. Deore, A. Pravin, Ensembling: model of histogram of oriented gradient based handwritten devanagari character recognition system. Traitement Du Signal 34(1–2), 7 (2017)

10. K.K. Kumar, H. Natraj, T.P. Jacob, Motion activated security camera using Raspberry Pi, in 2017 International Conference on Communication and Signal Processing (ICCSP) (IEEE, 2017), pp. 1598–1601

11. K. Madhu, R.I. Minu, Image segmentation using improved JSEG, in *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering* (IEEE, 2013, February). pp. 37–42

12. M. Sharma, A.K. Shaw, Transmission time and throughput analysis of EEELEACH, LEACH and direct transmission protocol: a simulation based approach. Adv. Comput. **3**(6), 75 (2012)

13. Meenu, Vandana, Modified PEGASIS in WSN to increase Network Lifetime. Int. J. Comput. Appl. **52**(19), 0975–8887 (2012)

14. S. Vishakha, S. Suri, Comparative study of hierarchical routing protocols in wireless sensor networks. Int. J. Comput. Sci. Eng. 2(5), 142–147 (2014)

15. R. Nathiya, S.G. Santhi, Energy efficient routing with mobile collector in wireless sensor networks (WSNs). Int. J. Comput. Sci. Eng. **2**, 36–43 (2014)

16. W. Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy-efficient communication protocols for wireless microsensor networks, in *Proceedings of the 33rd Hawaaian International*

Conference on Systems Science (HICSS), January 2000

17. S. Lindsey, C. Raghavendra, K.M. Sivalingam, Data gathering algorithms in sensor networks using the energy metric. IEEE Trans. Parallel Distrib. Syst. **13**(9), 924–935 (2002)

18. S. Vhatkar, M. Atique, Design Issues, characteristics and challenges in routing protocols for wireless sensor networks. Int. J. Comput. Appl. Int. Conf. Workshop Emerg. Trends Technol. (2013) pp. 0975–8887

19. S. Vhatkar, J. Rana, M. Atique, Performance Evaluation and QOS Analysis of EEPB and PDCH Routing Protocols in Wireless Sensor Networks. IOSR J. Comput. Eng. (IOSR-JCE) **17**(5), 101–109 (2015). ISSN: 2278–0661, p-ISSN:2278–8727

20. R. Mahakud, S. Rath, M. Samantaray, B. Sinha, P. Priya, A. Nayak, A. Kumari, Energy management in wireless sensor network using PEGASIS. Procedia Comput. Sci. **92,** 207–212 (2016)

21. P. Singh, S. Singh, A review: power efficient gathering in sensor information system (PEGASIS) protocol for WSN. Int. Res. J. Eng.

Technol. **2**(03) (2015). e- ISSN:2395–0056,p-ISSN:2395–0072.

22. D.G. Anand, H.G. Chandrakanth, M.N. Giriprasad, Energy-efficient clustering for Wireless Sensor Networks: a survey. Int. J. Adv. Res. Comput. Sci. (Case study and Report) **2**(6) (2011)

23. K. Chakravathy, U. Burman, An energy balanced algorithm of PEGASIS protocol in WSN. Int. J. Comput. Sci. Eng. **4**(7) (2016)

24. W. Jing, G. Yu, Y. Xisang, F. Li, J.K. Hye, An enhanced PEGASIS algorithm with Mobile Sink Support for Wireless Sensor Networks, in *Wireless Communication and Mobile Computing*, vol. 2018, Article ID 9472075 (2018)

25. Akyildiz et al., A survey on sensor networks. IEEE Commun. Mag. **40**(8), 102–144 (2002)

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

S. Sreeram & G. Nagarajan

Corresponding author

Correspondence to S. Sreeram .

Editor information

Editors and Affiliations

Birsa Institute Of Technology, Ranchi, Jharkhand, India
Neeraj Priyadarshi

Department of Energy Technology, Aalborg University, Esbjerg, Denmark
Sanjeevikumar Padmanaban

Department of Mechanical Engineering, Sikkim Manipal Institute of Technology, Majitar, Sikkim, India
Ranjan Kumar Ghadai

School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India
Amiya Ranjan Panda

School of Electrical Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India
Ranjeeta Patel

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Sreeram, S., Nagarajan, G. (2021). EDA-PEGASIS: A Balanced Energy Aware Routing Approach for Sensor Network to Reduce Cognitive Networking Complexities in Wireless Medium. In: Priyadarshi, N., Padmanaban, S., Ghadai, R.K., Panda, A.R., Patel, R. (eds) Advances in Power Systems and Energy Management. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 690. Springer, Singapore. https://doi.org/10.1007/978-981-15-7504-4_57

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7504-4_57

Published	Publisher Name	Print ISBN
21 January 2021	Springer, Singapore	978-981-15-7503-7

Online ISBN	eBook Packages
978-981-15-7504-4	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Double Door Authentication for Mobile Devices using Personalised Lock (pins)

[1]Chamarthi Tharun, [2]Chandra Rithin, [3]B.Bharathi
[1,2]Student, [3]Professor
[1,2,3]Department of Computer Science And Engineering
Sathyabama Institute of Science and Technology, Deemed to be University, Chennai, India
[1]ch.tharunkumar@gmail.com,[2]chandrarithin226@gmail.com, [3]bharathi.cse@sathyabama.ac.in

***Abstract*---As the usage of mobile devices has been enormously increased, the role of authentication became very prominent. Usually explicit authentication techniques such as passwords are employed. However the devices would be infracted if the passwords are stolen without the user's knowledge. Therefore, there is an adverse search for finding out the ways for improvising the authentication techniques for mobile platforms and services. Double door authentication for mobile devices is the method that provides privacy with enhanced security feature. This paper explores the main objective of providing an authentication facility that supports the users to manage all the confidential information with the feature of personalized locking(2 pins, hand-in-hand simultaneously).In a realistic manner, the personalized locking feature can exclude the devices from third-party security installations and interferences such as sure-lock, vault-hide etc. Here the primary and secondary users are considered to be the key players. Especially, the primary user has an extensive role of sensitive applications selection and restricted access implementation.**

***Keywords-- Mobile Devices, Authentication, Personalized Lock, Security, Breach, Explicit Authentication*.**

I. INTRODUCTION

The usage of mobile devices are frequent and extensive in our daily life[1]. With the help of improving technology, mobile devices are used as the medium for storage of sensitive and personal data. (Example: Banking and Identification information etc).Thus, the need of privacy is substantial in mobile devices and users are exactly looking for the same. According to the recent reports released by cybercrime and its respective organizations (2019), 80% of the mobile users complain with the concern of sensitive information leakage and the rest face the problem of unknown content loss from mobile devices [2]. As the need for privacy is increased, it is necessary to develop a system where the sensitive information can well be protected and secured before loss. Well-known and identified method of user authentication are Passwords. Still many users do not know the ways of affixing a strong passwords for device protection. Also, many users tend to forget the passwords and to resolve this problem some few authentication features were developed by the mobile developers. The techniques are finger-print scanner [3], face recognition, speech authentication, and Android face-unlock etc. The above mentioned are only based on the human physiological characteristics and to better improvise the system something great can be implemented depending on the environmental conditions.

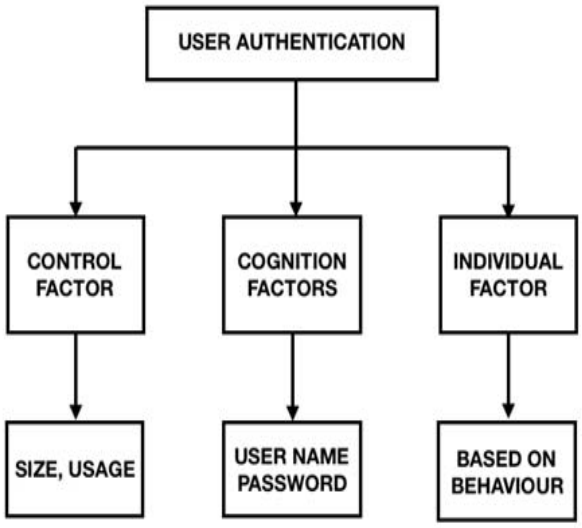


Fig.1 Factors of Mobile Authentication System.

Authentication can be classified based on authentication factors. The authentication can be of Single-Factor (1FA) and Two-Factor (2FA).In case of single factor authentication only one authentication procedure is employed. But, in two-factor authentication sum of authentication procedures are involved.(Eg: E-Banking where both password and OTP are essential).This paper provide with the theme of controlled access for secondary users with the support of personalized pin thereby neglecting the additional installations from mobile play-store[4]. Thus, personalized pin can prevent the loss of information and devices can be managed securely. Hence, it is the primary user/admin who needs to decide the selection of sensitive applications for restricted access to the secondary user.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.



II. LITERATURE SURVEY

A) Authors Amy K.Karlson and A.J. Bernhein Brush in the paper about "Can I Borrow Your Phone?"(2015)-Understanding problems when mobile devices are shared, studied and emphasized the multifariousness of guest user classification and related security reasons that was stated by the users and suggested security model better than binary model. The users mainly focuses on the theme of guest user logins. Though, guest users are temporary but the problem faced by those users are much more in number. The guest users could also face serious accessibility issues and the drawbacks need to be resolved and the efficiency of the system can be improved for privacy reasons.

B) In the paper "A Survey Of Biometric Keystroke Dynamics(2016): Approaches, security and challenges, the authors D.Shanmugapriya and G.Padmavathi emphasized on the process of keystroke Dynamics. Keystroke Dynamics is a prominent technology of Biometrics. It will exactly identify the authentication of user when the user is working via keyboard accessory. The authors discussed various key security issues and challenges that are faced with Keystroke Dynamics. Accuracy of the system was enhanced using keystrokes and in cases of inefficiency, it will lead to various challenges. Keyboard plays a very prominent role where all the necessary entries are made with the keyboard. It also depends on the number of entries.

C) Authors Mr.Ka-Wing Tsee,Kevin-Hung in the paper "Swipe Dynamics For User Authentication On Mobile Platforms"(2017) which studied and focused on authentication schemes that combines the passwords, keystroke dynamics and swipe dynamics. It generally deals with the touch screen mobile devices. Here the authentication can be deployed in multiple ways. The authors evaluation is based on swipping patterns and typing patterns. Using the above mentioned techniques, accuracy of the system was enhanced particularly with behavioural Biometric features. Thus, the identification accuracy grew extensively to wider ranges upto 78%.

D) In the paper "Continuous Authentication Of Smartphones Based On Application Usage"(2018) authors namely Upal Mahbub, Jukka Komulainen and Denzil Ferreira emphasized on the concept of active authentication where the verification is based on the usage time. Here the users made the use of Markov Models, Marginal smoothing and Sequence Matching. It is mainly based on the theme of Application Usage. For Example, when there are two applications say Application 'A' and Application 'B', if the usage of application 'A' is more frequent than Application 'B', it clearly indicates the greater usage activity of Application 'A' than for Application 'B.' This way the authors made a fantastic contribution of prioritization and grouping of several applications. The implementations based on priority shall be easy and worthwhile.

E) Authors Tempestt Neal and Damon Woodard in the paper "You Are Not Acting Like Yourself"(2018), which studied about the Biometric Classification, Personal Identification and Mobile devices Use. Here the authors emphasized on the classification of thirteen soft biometric classes with the help of calling, text messages and application data. The usefulness has been clearly known and determined. Authors also defined the combination and distribution of data for various statistical reasons. The ways of implementation also increases the efficiency of the system in-place.

III. EXISTING SYSTEM

Mobile Devices are protected and secured with the aid of individual locking mechanism. The locking could be worthwhile in case of single usage and at times when the device is shared with the secondary user, the secondary users can still have usage to all the applications that are been accessed by the admin[2]. This system still lacks in privacy and security to the mobile admin. Besides, the system can still lead to chances of information violation though the secondary user is trusted.

Apart from the in-built features, the mobile users even had a choice of downloading and installing the third-party applications such as Gallery Vault, Secured Folder Apps etc for additional advantage[5]. This seems to be appropriate but not fulfilling to the user. It takes much of the resources to be additionally spent and experienced. For a user, advanced privacy techniques can have better utility in mobile usages in terms of integrity and confidentiality. The users cannot completely rely on the single/multi-existing normal locks and some sort of personalised feature can make the mobile system to be more efficient in security and performance related concerns.[6]

IV. PROPOSED SYSTEM

The proposed system focuses on the assistance of personalized locking where the users utilise the concept of improving mobile privacy by adding a personalized pin for unlocking the device. Here the deployment of two-pin conception comes into existence. Generally

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.

mobile devices can be accessed by the admin/primary user (or) the secondary user. When the primary user intends to access the mobile devices with the device's original pin **(eg.1234)**[12], the users are permitted to access all the applications of the mobile phone where the primary user makes a clear decision regarding the sensitive applications and selects those applications which has to be hidden from the secondary user. Now, when the secondary users access the mobiles with the personalized pin **(eg:5678)**[13], secondary users can only access the normal applications available and cannot have an eye on the sensitive/personal applications as the admin will make them to be under hidden as stated.

The advantage of the personalized lock is that the secondary users cannot even know the additional feature that has been binded in the device and no question of talk arises from the secondary user[11]. The concept has been employed keeping in mind all the necessary advantages of the Mobile Devices. So, the Primary user (admin) has a greater reliable privacy when the device is physically shared with the secondary user at times for casual purposes[2]. As the secondary user could not know the feature of locking, it supports the mobile devices to be much safer and no chances of breaching occurs. This system provides greater utility to the users along with the theme of confidentiality and integrity.

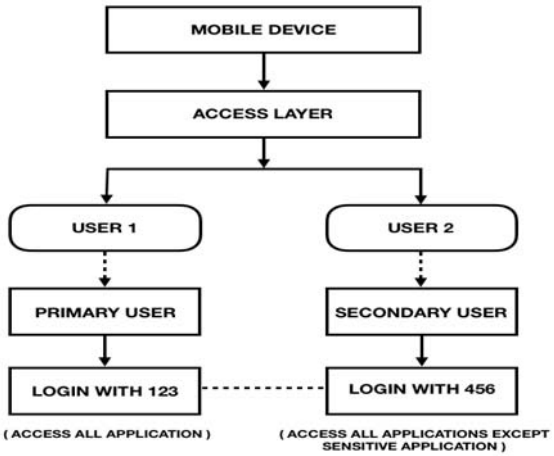


Fig.3 Flow of the application

EXAMPLE: Many today wanted to make many things to be personal on the device. Though the secondary user is a trusted one, all the information cannot be shared and the admin sometimes expects privacy. The applications such as What's App, Instagram, Facebook and all other social media applications[7] especially, Banking Applications lack in security few times. Though the security is maintained, it is not upto the mark of utility. So, the proposed system can have wider range of usage in real time scenario's.

V. METHODOLOGIES

A) MOBILE APPICATION DEPLOYMENT

Mobile client is an android application which is created and installed in the user's android device to perform the activities. The usage of android emulator plays a prominent role in physical implementation. The Application First Page consists of the user integration process. Here, the creation of the Use Login Page and text field by Button is necessary. When the mobile application is generated, it is stored as APK (Application Package) for execution on the android platform. The APK files will be installed in the user's mobile phone as an Application. The emulator clearly represents the activities performed.

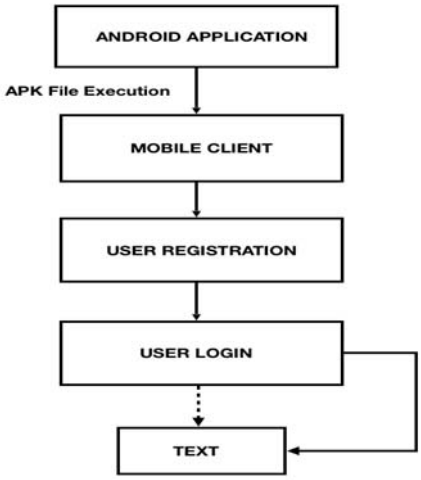


Fig.4. Representation of the Android Application Execution

B) USER-ACCESSIBILITY DEPENDING ON USER-PERMISSION

As the primary users are considered as the admins of the mobile devices, the establishment of a password to hide the sensitive applications from restricted access is deployed. Secondary users can only view the set of applications. The primary advantage of the process is that the primary user can restrict the complete accessibility to the Secondary user with the aid of personalized feature.

C) DATA COLLECTION

For the processes to prolong, the data for application need to be collected. Data needs to be collected in such a manner where the data needs to be in a sequential and understandable format. The kind of data collected includes the features of various domains and classification is thereby performed based on the utility.

D) POLICY-PRIVACY SETTING FOR SENSITIVE MOBILE APPLICATIONS

The sensitive applications need to be clearly specified so that the authority will be restricted for secondary users. The Admin are provided with application

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.

selection feature where the new user cannot be able to view other applications. The feature of privacy setting for sensitive application is established and security is assigned to those applications.

E) DEPLOYMENT OF DUAL PASSWORD(PIN/PATTERN)

It includes the method of setting up a password either in the form of secured pin/pattern according to the user's convenience. It deals with the password setup depending on the priority of applications. It also depends on the usage of the devices. Security locking mechanism is implemented and the setup could be the right choice for privacy.

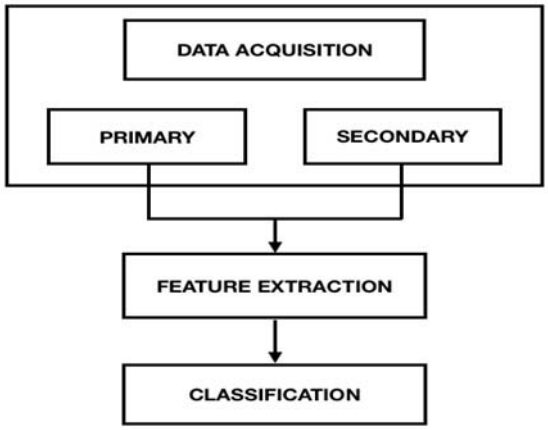


Fig.5. Representation of Activity (ON state)

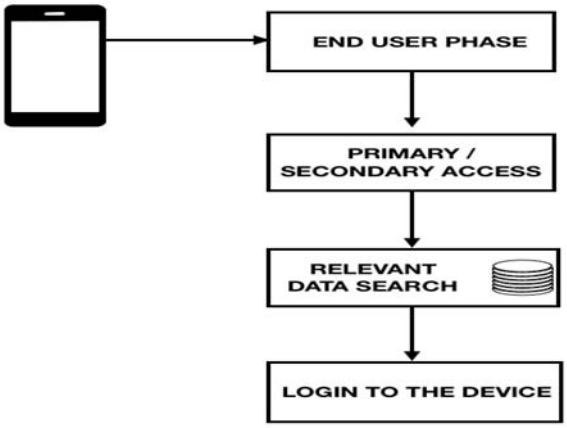


Fig 6. Deployment of Dual password through sequential search

VI. RESULTS AND EVALUATION

Double door authentication for mobile devices could be an additional personalized feature for mobile devices. It provides the users with greater security performance for protection of sensitive data in mobile devices by setting up of the personalized pins on the user's desk.



Fig.7 Representation of the Registration.



Fig. 8: Off State

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.



Fig. 9: On state



Fig 10. Primary User



Fig. 11: Secondary user

Sequential procedure where state mode, registration of users and pin setup has been defined for the implementation. This can enable the personalized locking technique of highly secured information and can worthwhile be prevented and protected from information loss. Many a times, it is a situation for users to classify the application privacy based on prior priority. The simultaneous usage of both the normal pin and personalized pin can be useful w.r.t user in place. However, analyzing both the positive and negative aspects, it could be satisfactory for any user to use the feature of personalized locking. So, the concept of Double Door Authentication for Mobile Devices will make the mobile usage more efficient.

VII. FUTURE ENHANCEMENT AND CONCLUSION

In the up-coming years, the personalized feature can directly be implemented in the “Operating System Kernel”. As the personalized locking mechanism is not known to secondary user, it can have a dwelling advantage for the primary user’s. In the futher process, it can also be implemented with the help of advanced Facial Recognition Feature. As, the secondary users cannot identify the mobile user accounts/passwords/access securities and other device’s data, it is beneficial in implementation. In the future part of work, it is recommended to deploy the implementation in the O.S kernel where the system can work from the back-end through certain operating system modifications. This mechanism can eventually have a greater role apart from other locking systems and

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.

the user can completely ignore the third-party installations/locking applications. It is adequate for the user to make use of the in-built lock feature of the mobile devices.

REFERENCES

1. Banerjee, S., & Woodard,(2016) D., "Biometric authentication and identification using keystroke dynamics: a survey," Journal of Pattern Recognition Research, 7(1), 116–139, 2016.
2. Giot, Romain,(2015) El-Abed, Mohamad& Rosenberger, Christophe, "Keystroke Dynamics Authentication," 2015.
3. G. Forsen, M. Nelson, R. Staron, Jr.,(2015) "Personal attributes authentication techniques", Technical Report RADC-TR-77-333, Rome Air Development Center, October 2015.
4. Karlson, A.K., Brush, A.B., Schechter. S,(2015),"Can I Borrow Your Phone?: Understanding Concerns When Sharing Mobile Phones," Proceedings of the 27th International Conference on Human Factors in Computing Systems (CHI), pp. 1647–1650, ACM, New York, USA, 2015.
5. Number of mobile phone users worldwide from 2015 to 2020 (in billions). Retrieved from the provided https://www.statista.com /statistics/274774/forecast-of-mobile-phoneusers-worldwide/.
6. Mr. Ka-Wing TSEE and Kelvin-Hung (2017) focused on "Swipe Dynamics for user authentication on mobile platforms, which studied and focused on authentication schemes that combines the passwords, keystroke dynamics and swipe dynamics.
7. R. Spillane,(2015) "Keyboard Apparatus for Personal Identification," IBM Technical Disclosure Bulletin, vol. 17, no. 3346, 2015.
8. Shanmugapriya, D., &Padmavathi, (2016) G., "A survey of biometric keystroke dynamics: approaches, security and challenges," International Journal of Computer Science and Information Security, 5, 115–119, 2016.
9. S. Bleha, C. Slivinsky, and B. Hussien,(2015) "Computer-access security systems using keystroke dynamics," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 12, pp. 1217–1222, 2015.
10. Singha, A.K., Singla and Pandey, R. K., (2016) "Study and analysis on biometrics and face recognition methods," EPH-International Journal of Science And Engineering (ISSN: 2454-2016), 2(6):37–41, 2016.
11. Tempestt Neal and Damon Woodard, (2018) "You are not Acting like Yourself", which studies about the Biometric Classification, Personal Identification and Mobile devices Use.
12. Upalmahbub, jukkakommulainen and denzil ferreira,(2018)"Continuous Authentication of smart phones based on application usage, emphasized on the concept of active authentication where the verification is based on the usage time. Here the users made the use of markov models, marginal smoothing and sequence matching.
13. Wayman, J., Jain, A., Maltoni, D., and Maio, D.(2015), "An introduction to biometric authentication systems," Biometric Systems, pages 1–20, Springer, 2015.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:25:20 UTC from IEEE Xplore. Restrictions apply.

Advertisement

WILEY
KnowItAll Campus So
Chemistry and Spectral Resources

Angewandte International Edition Chemie

GDCh A Journal of the German Chemical Society

Research Article

Discovery of a Novel Mycobacterial F-ATP Synthase Inhibitor and its Potency in Combination with Diarylquinolines

Adam Hotra, Priya Ragunathan, Pearly Shuyi Ng, Pattarakiat Seankongsuk, Amaravadhi Harikishore, Jickky Palmae Sarathy, Wuan-Geok Saw, Umayal Lakshmanan, Patcharaporn Sae-Lao ... See all authors

First published: 26 April 2020
https://doi.org/10.1002/anie.202002546
Citations: 18

Get access to the full version of this article. View access options below.

Institutional Login

Sathyabama Dental College & Hospita does not provide access to this content.

Log in with a different institution

Log in to Wiley Online Library

If you have previously obtained access with your personal account, please log in.

Log in

Purchase Instant Access

48-Hour online access $15.00

Details

Online-only access $25.00
Details
PDF download and online access $59.00
Details
Check out

Abstract

The F_1F_O-ATP synthase is required for growth and viability of *Mycobacterium tuberculosis* and is a validated clinical target. A mycobacterium-specific loop of the enzyme's rotary γ subunit plays a role in the coupling of ATP synthesis within the enzyme complex. We report the discovery of a novel antimycobacterial, termed GaMF1, that targets this γ subunit loop. Biochemical and NMR studies show that GaMF1 inhibits ATP synthase activity by binding to the loop. GaMF1 is bactericidal and is active against multidrug- as well as bedaquiline-resistant strains. Chemistry efforts on the scaffold revealed a dynamic structure activity relationship and delivered analogues with nanomolar potencies. Combining GaMF1 with bedaquiline or novel diarylquinoline analogues showed potentiation without inducing genotoxicity or phenotypic changes in a human embryonic stem cell reporter assay. These results suggest that GaMF1 presents an attractive lead for the discovery of a novel class of anti-tuberculosis F-ATP synthase inhibitors.

Conflict of interest

G.G., R.W.B., A.H., T.D., K.P. are inventors on the patent PCT/SG2018/050075, and G.G., R.W.B., P.S.N. are inventors on the patent 10201911205R which are related to the chemical series described in this article.

Supporting Information

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Filename	Description
anie202002546-sup-0001-misc_information.pdf 9.2 MB	Supplementary

Please note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.

Download PDF

About Wiley Online Library

Privacy Policy
Terms of Use
About Cookies
Manage Cookies
Accessibility
Wiley Research DE&I Statement and Publishing Policies
Developing World Access

Help & Support

Contact Us
Training and Support
DMCA & Reporting Piracy

Opportunities

Subscription Agents
Advertisers & Corporate Partners

Connect with Wiley

The Wiley Network
Wiley Press Room

Copyright © 1999-2023 John Wiley & Sons, Inc. All rights reserved

WILEY

A Journal of the Gesellschaft Deutscher Chemiker
Angewandte
GDCh
International Edition
Chemie
www.angewandte.org

Accepted Article

Title: Discovery of a novel Mycobacterial F-ATP synthase inhibitor and its potency in combination with diarylquinolines

Authors: Adam Hotra, Priya Ragunathan, Pearly Shuyi Ng, Pattarakiat Seankongsuk, Amaravadhi Harikishore, Jickky Palmae Sarathy, Saw Wuan Geok, Umayal Lakshmanan, Patcharaporn Sae-Lao, Nitin Pal Kalia, Joon Shin, Revathy Kalyanasundaram, Sivaraj Anbarasu, Krupakar Parthasarathy, Chaudhari Namrata Pradeep, Harshyaa Makhija, Peter Dröge, Anders Poulsen, Jocelyn Hui Ling Tan, Kevin Pethe, Thomas Dick, Roderick W. Bates, and Gerhard Grüber

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202002546

Link to VoR: https://doi.org/10.1002/anie.202002546

WILEY-VCH

10.1002/anie.202002546

RESEARCH ARTICLE

Discovery of a novel Mycobacterial F-ATP synthase inhibitor and its potency in combination with diarylquinolines

Adam Hotra[1,2,3], Priya Ragunathan[2], Pearly Shuyi Ng[4], Pattarakiat Seankongsuk[1], Amaravadhi Harikishore[1,2], Jickky Palmae Sarathy[5], Wuan-Geok Saw[2], Umayal Lakshmanan[4], Patcharaporn Sae-Lao[1], Nitin Pal Kalia[6], Joon Shin[2], Revathy Kalyanasundaram[7], Sivaraj Anbarasu[7], Krupakar Parthasarathy[7], Chaudhari Namrata Pradeep[2], Harshyaa Makhija[2], Peter Dröge[2], Anders Poulsen[4], Jocelyn Hui Ling Tan[4], Kevin Pethe[2,6], Thomas Dick[5,8,9], Roderick W. Bates[1,*], and Gerhard Grüber[2,*]

Abstract: The F_1F_O-ATP synthase is required for growth and viability of *Mycobacterium tuberculosis*. Bedaquiline validated this target clinically. We showed recently that a mycobacterium-specific loop of the enzyme's rotary γ subunit plays a role in coupling of ATP synthesis within the enzyme complex. Here, we report the discovery of a novel antimycobacterial, termed GaMF1, targeting this γ subunit loop. Biochemical and Nuclear Magnetic Resonance studies show that GaMF1 inhibits ATP synthase activity by binding to the loop. GaMF1 is bactericidal and is active against multidrug- as well as Bedaquiline-resistant strains. Chemistry efforts on the scaffold revealed a dynamic structure activity relationship and delivered analogues with nanomolar potencies. Combining GaMF1 with bedaquiline or novel diarylquinoline analogues showed potentiation without inducing genotoxicity or phenotypic changes in a human embryonic stem cell reporter assay. These results suggest that GaMF1 presents an attractive lead for the discovery of a novel class of anti-tuberculosis F-ATP synthase inhibitors.

[1]School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Republic of Singapore; E-mail: roderick@ntu.edu.sg
[2]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Republic of Singapore; E-mail: ggrueber@ntu.edu.sg
[3]Nanyang Institute of Technology in Health and Medicine, Interdisciplinary Graduate School, Nanyang Technological University, Republic of Singapore
[4]Experimental Drug Development Centre, Agency for Science Technology and Research, A*STAR, 10 Biopolis Road, Singapore 138670
[5]Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, 14 Medical Drive, Singapore 117599, Republic of Singapore
[6]Lee Kong Chian School of Medicine, Nanyang Technological University, Experimental Medicine Building, Singapore
[7]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai 600119, India
[8]Center for Discovery and Innovation, Hackensack Meridian Health, 340 Kingsland Street, Nutley, NJ 07110, USA
[9]Center for Discovery and Innovation, Hackensack Meridian Health, 340 Kingsland Street, Nutley, NJ 07110, USA

Introduction

Tuberculosis (TB) is a devastating disease that results in 1.4 million deaths every year.[1] Multidrug resistant- (MDR) and extensively drug-resistant (XDR) bacterial strains of *Mycobacterium tuberculosis* (*Mtb*), which causes TB, are spreading worldwide, causing major global health issues. Furthermore, about one third of the world population is latently infected by Koch's bacillus and susceptible to develop the disease.[1] During infection, *Mtb* inhabits a wide range of intracellular and extracellular environments. Many of these environments are highly dynamic, and mycobacteria face with the constant challenge of redirecting their metabolic activity commensurate with either growth or non-replicating persistence.[2] A fundamental feature in this adaptation is the ability of mycobacteria to respire, regenerate reducing equivalents, and generate the biological energy currency adenosine triphosphate (ATP) via oxidative phosphorylation. During the latter, an electrochemical gradient of protons is generated by the electron transport chain (ETC; Fig. 1A). The electrochemical gradient drives the energy converter F_1F_O ATP synthase and is required for viability of growing and non-growing mycobacteria.[2] Phenotypic screening identified the first-in-class diarylquinoline drug bedaquiline (BDQ, Sirturo®) that inhibits the mycobacterial F-ATP synthase (Fig. 1A-B).[3] However, the successful advance of this drug has been overshadowed by the development of clinical resistance [4] along with a very long terminal half-life leading to concerns regarding tissue accumulation.[5] Recent medicinal chemistry campaigns resulted in a new generation of 3,5-dialkoxypyridine (DARQ) analogues of BDQ that have the potential to address these issues.[6] TBAJ-587 and -876 are new development compounds of this series (Fig. 1B), showing improved physio-chemical properties.[6] The discovery of squaramides (SQA) [7], compound 5228485 [8], and epigallocatechin gallate (EGCG) [9] as potent ATP synthesis inhibitors underlined the relevance of the mycobacterial F-ATP synthase as a drug target.

Nevertheless, infections with MDR- and XDR *Mtb* strains and coinfection with HIV have worsened the prospect of controlling TB.[10] In addition, the prolonged treatment regimen for six months with a combination of four drugs has resulted in poor patient compliance, leading to the emergence of MDR and XDR strains of *Mtb.*[11] Therefore, in order to improve the effectiveness of TB treatment and to improve control of MDR and XDR TB, new antitubercular targets and agents are required. Anti-TB targets should be a) essential for viability in *Mtb*, b) novel to avoid of cross resistance with existing drugs, c) absent in humans to avoid mechanism-based toxicity issues.

The F_1F_O ATP synthase is essential for growth and viability in mycobacteria.[12] In contrast, in other prokaryotes, the enzyme is dispensable for growth on fermentable carbon sources.[13] The mycobacterial enzyme is composed of a membrane-embedded F_O complex (subunits a:b:b':c_9), responsible for H^+-translocation from the periplasmic space to the cytoplasmic side, and the F_1 part (subunits α_3:β_3:γ:ε) [14] which uses the proton-motive force to synthesize ATP within the three catalytic αβ-pairs (Fig. 1A). This catalytic $\alpha_3\beta_3$-headpiece is linked with the F_O part via the two rotating central stalk subunits γ and ε, as well as the peripheral stalk subunits b, b' and δ.[14c, 14e].

We recently demonstrated the specific role of the unique loop of mycobacterial subunit γ (γ166-179-loop; *M. smegmatis*

This article is protected by copyright. All rights reserved.

Accepted Manuscript

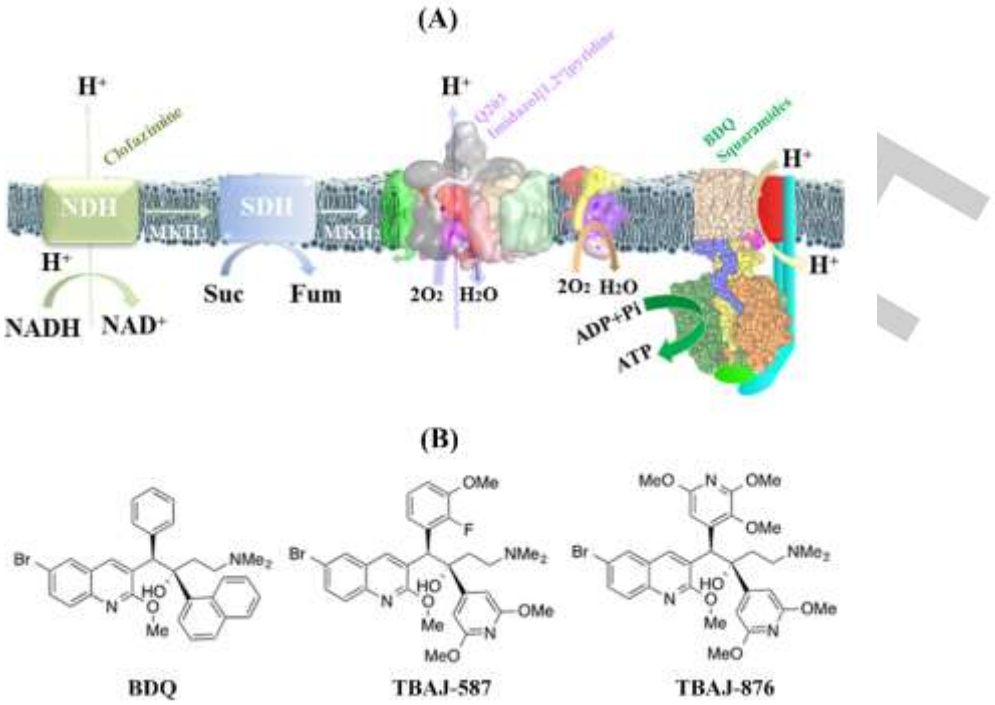


Figure 1. (*A*) Enzyme composition and inhibitors of the electron transport chain and F-ATP synthase of *Mtb*. (*B*) Structures of BDQ, TBAJ-587 and TBAJ-876.

numbering) effecting ATP hydrolysis, -driven H^+-pumping and -synthesis.[15] As the mycobacterial γ-loop is unique, absent in the human homologue and other prokaryotes, as well as important for the regulation of ATP synthesis, -hydrolysis and proton pumping, it fulfils fundamental criteria as a potential TB-drug target.

We here conducted an *in silico* compound screening exercise based on our structural and biochemical data of the mycobacterial enzyme F-ATP synthase [15, 16] that led to the discovery of the mycobacterial F-ATP synthase inhibitor GaMF1. The variety of biochemical assays, the use of a BDQ-resistant F-ATP synthase mutant and structural experiments validated that GaMF1 exerts its anti-mycobacterial activity through inhibition of the F-ATP synthase. Furthermore, GaMF1 increased the potency of BDQ and other DARQ analogues, opening a new avenue for an efficient multi-drug combination. In addition, initial structure-activity relationship studies (SAR) revealed a path to optimize the potency of the series.

Results and Discussion

Homology and Pharmacophore Modelling

The γ166-179 loop was exploited as a binding epitope for *de novo* drug design. [15, 16] As high-resolution crystallographic structures were lacking when the project started, homology

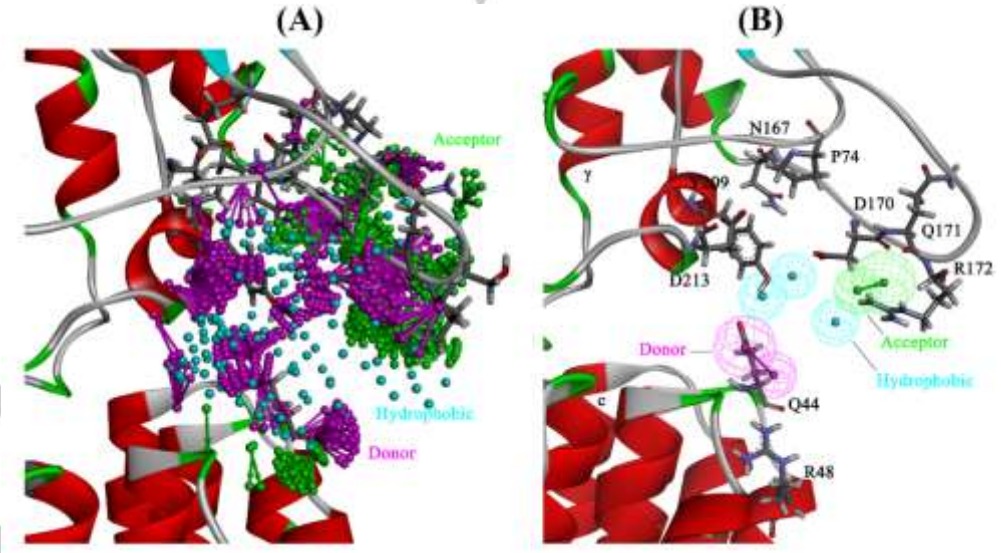


Figure 2. Receptor-based pharmacophore modeling and pharmacophore mapping. (*A*) Generation of Ludi-based interaction sites using acceptor, donor, and hydrophobic features as probes at γ166-179 loop-*c* ring interface residues. (*B*) Five features comprising of one hydrogen bond acceptor (*green sphered arrows*) in vicinity of R172 sidechain atoms, one hydrogen bond donor targeting Q44 residue of the c-ring (*magenta sphered arrows*), three hydrophobic features (*cyan spheres*) in the vicinity of P74, Y209, R172 residues were included in the receptor based pharmacophore model.

This article is protected by copyright. All rights reserved.

Accepted Manuscript

modelling enabled us to build 3D structures with good accuracy (see Supplementary Information). Our receptor-based pharmacophore modelling studies highlighted the possible interaction sites at the γ166-179 loop interface (Fig. 2A). Mainly, one hydrogen bond acceptor, one hydrogen bond donor and three hydrophobic features were included in the pharmacophore model at the γ166-179 loop-*c* ring interface (Fig. 2B). Further, the exclusion volumes constraints around protein atoms near the ligand (not shown for clarity) were also incorporated to delineate the shape of the binding cavity as well as to prevent clashes of ligand with protein atoms. Database screening with a 5-feature mapping resulted in library of about 30,000 molecules. This library was further refined by a series of standard precision, extra precision docking protocols with glide score, XP scores and consensus scoring themes. Together with visual inspection of molecular interactions at the γ166-179 loop and *c*-ring interface resulted in 81 molecules for experimental screening.

Antimycobacterial activity and Minimum inhibitory concentration determination

We identified ligands that bear complementarity to the γ166-179 loop and *c*-ring interface and provided a focused library of which GaMF1 (Fig. 3A) exhibited a minimal inhibitory concentration (MIC_{50}) of 11 µM in the fast growing *M. smegmatis* mc^2 155 strain, 17 µM in *M. bovis* bacillus Calmette–Guérin (BCG), and 33 µM in *Mtb* H37Rv (Fig. 3B). A synthesis protocol for GaMF1 was established (see Supplementary Information: *Synthesis of GaMF1*) to produce the compound in bulk to study the bactericidal activity of the compound and its mechanism of action. GaMF1 was bactericidal against *M. smegmatis* mc^2 155 at 20-fold its MIC_{50}, indicated by the decreasing cell number of about 99% in 3 days (Fig. 3C) In contrast, BDQ had delayed bactericidal activity, as reported before [17], followed by a reduction in bacterial viability (Fig. 3C). GaMF1 was also bactericidal

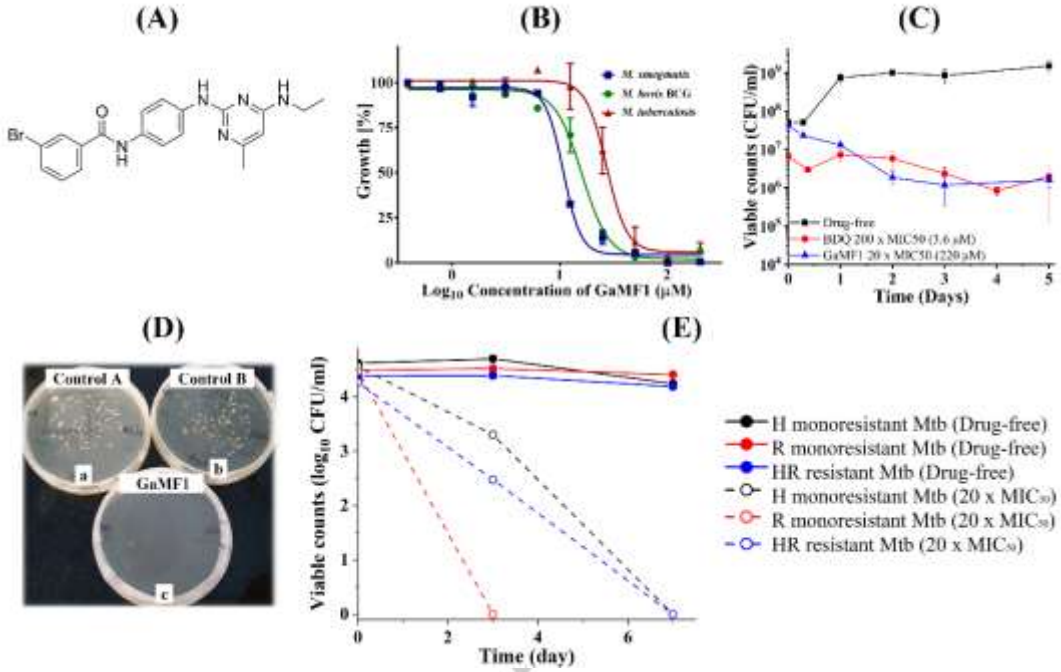


Figure 3: (*A*) Structure of GaMF1. (*B*) Growth inhibition of *M. smegmatis* (*blue*), *M bovis* BCG (*green*) and *Mtb* (*red*) by GaMF1 with MIC_{50} of 11, 17 and 33 µM, respectively. (*C*) Initial five days of GaMF1-, BDQ kill kinetics against *M. smegmatis* mc^2 155. The bacteria were grown in liquid culture (LBT) in the presence of the indicated concentrations of GaMF1 and BDQ. The experiments were repeated twice, and the profiles were identical. (*D*) Images (*a*, *b*), HR resistant *Mtb* colonies observed after day 7 incubation in no-compound containing Middlebrook 7H9 broth (4.188 log_{10} CFU/ml); Image (*c*), no colonies of HR resistant *Mtb* observed after day 7 incubation in compound GaMF1 containing Middlebrook 7H9 broth. (*E*) Antimycobacterial activity of GaMF1 against *Mtb* isolate resistant to Isoniazid (H) and Rifampicin (R) as determined by Colony forming units (CFU) based method in Middlebrook 7H11 agar plates.

against three drug resistant clinical isolates (isoniazid (H), rifampicin (R) monoresistant- and HR-resistant *Mtb* strains), causing a significant reduction in bacterial viability in 3 to 7 days (Fig. 3D- E). Preliminary *in vitro* ADME profiling revealed that that GaMF1, with a clogP value of 4.37, is less lipophilic than BDQ [6c], and has a good metabolic stability in mouse liver microsomes ($T_{½}$ with or without NADPH of 47.5 min and 68 min, respectively; Cl_{int} of 14.6 µl/min/mg protein (Supplementary Table 1A-B)).

Insights into the inhibition of mycobacterial F-ATP synthase by GaMF1

To confirm that GaMF1 exerts its anti-mycobacterial activity via inhibition of the mycobacterial F-ATP synthase, ATP synthesis was measured using inverted membrane vesicles (IMVs) of *M. smegmatis* and *M. bovis* BCG. Figure 4A shows a typical profile of *M. smegmatis* IMVs ATP synthesis inhibition by BDQ with a half-maximal inhibitory concentration (IC_{50}) of 1.1 ± 0.2 nM, which is in line with previous reports [15, 18], and an IC_{50} of 7.1 ± 1.3 nM *M. bovis* BCG (Fig. 4B). In comparison, GaMF1 revealed a potent ATP synthesis inhibition at an IC_{50} of 0.5 ± 0.1 µM in *M. smegmatis* IMVs (Fig. 4A) and 5.2 ± 1.1 µM in *M. bovis* BCG (Fig. 4B). GaMF1 inhibited ATP synthesis of *M. smegmatis* (Fig. 4A) and *M. bovis* BCG (Fig. 4B) IMVs in the presence of succinate, indicating that GaMF1 does not interfere with the NADH-dehydrogenases. In parallel, a possible effect of GaMF1 on oxygen consumption and thereby the cytochrome oxidases of the ETC was tested using methylene blue as an oxygen probe. As visualized for the fast- and slow grower *M. smegmatis* (Fig. 4C) and *M. bovis* BCG (Fig. 4D), oxygen consumption was unaffected in both strains at µM concentrations of GaMF1. In comparison, nM- and µM concentrations of BDQ revealed a slight and moderate effect on oxygen respiration in both strains (Fig 4C-D), reflecting that BDQ does also affect oxygen consumption as described by Kalia et al. [19]

This article is protected by copyright. All rights reserved.

Accepted Manuscript

10.1002/anie.202002546

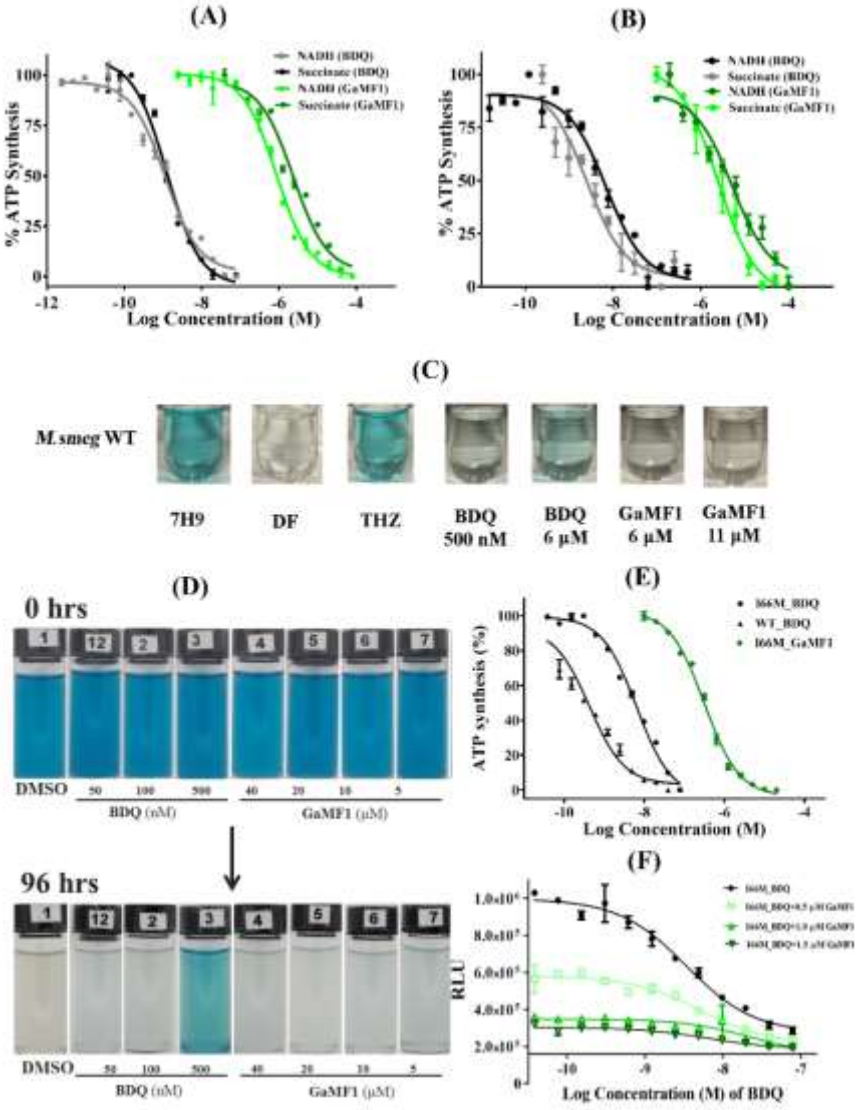


Figure 4: Inhibition of ATP synthesis by GaMF1 in IMVs from *M. smegmatis* (*A*) and *M. bovis* (*B*) using the electron donors NADH (*dark green*) or succinate (*light green*). BDQ was used as a control BDQ (*black/grey*). (*C-D*) Test of possible effects of GaMF1 and BDQ on oxygen consumption in *M. smegmatis* IMVs (*C*) and *M. bovis* BCG cells (*D*). BDQ shows partial inhibition of oxygen consumption at nanomolar range (500 nM) and significant inhibition at micro-molar level (6 µM). In comparison, GaMF1 did not inhibit oxygen consumption even at high micro-molar level (6- or 11 µM). Thioridazine (THZ) was used as positive control and 7H9 media /drug free (DF) served as a blank/negative control (*C*). GaMF1 did also not affect oxygen consumption in the slow grower *M. bovis* over a 96-h period, while BDQ at 500 nM completely inhibited oxygen consumption (*D*). (*E*) Inhibition of ATP synthesis by BDQ (*black*) and GaMF1 (*green*) on *M. smegmatis* I66M *c* subunit mutant IMVs with NADH as substrate. The *M. smegmatis* I66M mutant IMVs (●) reveal a 10-fold sensitivity compared to WT (▲), while the inhibitory profile of GaMF1 on *M. smegmatis* I66M mutant IMVs (♦) did not alter. (*F*) Additive effect of GaMF1 and BDQ (*black*) on the ATP synthesis inhibition in IMV's of *M. smegmatis* I66M mutant. In the I66M mutant, BDQ cannot fully inhibit the ATP synthesis. However, when GaMF1 (*green*) was added in a concentration dependent manner, ATP synthesis significantly decreased to almost full inhibition.

We next tested the potency of GaMF1 against an F-ATP synthase bearing an I66M substitution in the *M. smegmatis* subunit *c*, which is associated with resistance to BDQ. [20] While the mutant displayed a 10-fold shift in sensitivity to BDQ (Fig. 4E), the mutation did not affect the potency of GaMF1 (Fig. 4E). Interestingly, addition of GaMF1 reduced ATP synthesis of the I66M mutant IMVs significantly in a concentration dependent manner (Fig. 4F), demonstrating that the novel F-ATP synthase inhibitor GaMF1 shows no cross resistance to BDQ.

Target specificity of GaMF1

The *in vitro* assay data revealed that GaMF1 inhibits ATP synthesis and the growth of *M. smegmatis* and *M. bovis* BCG strains. Next, we performed NMR titration experiments to assess whether GaMF1 does bind to the *M. tuberculosis* γ166-179 peptide, which according to its NMR solution structure, forms a loop of polar residues with a surface distribution of negatively charged residues on one side and an amphiphilic charge distribution on the other side of the loop.[16] The chemical shift perturbations (CSPs) for NH and Hα between free peptide and the peptide-GaMF1 complex derived from the 2D-TOCSY spectra of the *Mt*γ166-179-loop with GaMF1 suggest that the amino acids N167, G168, E169, R172, and S173 became perturbed upon ligand binding (Fig. 5A-B). The data reflect the involvement of γ166-179-loop residues in GaMF1-binding.

To obtain further insights into the molecular interactions of GaMF1, we docked the ligand into subunit *Mt*γ using the NMR perturbed residues as ambiguous interaction restraints (AIR) or active site residues for ligand docking (Fig. 5C). GaMF1 comprises of a 3-bromo-benzoyl moiety (ring A), a 1,4 diaminobenzene moiety (ring B) and a substituted pyrimidine (ring C) and was predicted to bind to a region in the vicinity of the *Mt*γ-

This article is protected by copyright. All rights reserved.

Accepted Manuscript

loop residues with a glide score of 5.58 (and a GOLD PLP fitness of 46.5). Both ring A and -B of GaMF1 are involved in interactions with the γ166-179 loop residues. On one hand, the carbonyl and the amide atoms on ring A were engaged in close contacts with γQ171 side-chain amide atoms and γD64 main-chain amide atoms. On the other hand, the 3-bromo benzene group on ring A was also involved in π-alkyl interactions with amino acids γH65 and alkyl atoms of γ R172 (*highlighted by pink lines*). This moiety is positioned in the vicinity of the *c*-ring residues such as *c*R52, *c*Q46, *c*P47, *c*E48 (*beige coloured*) via close contacts, and thereby chaperone the stability of ligand binding at the *Mt*γ-*c* interface. The phenyl-ring (ring B) and pyrimidine (ring C) of GaMF1 were engaged in π-anion interactions (*orange arrows*) with γE169 and γE70 residues, respectively. The 6-methyl on ring C maintained aliphatic hydrophobic interactions with γI214 (*pink line*, (α2 / β5 loop, *light red*)). In addition, the amide atoms which link the 3-N-ethyl fragment on ring C mediate the hydrogen bonding interactions with γR71 residues from α1/β1 loop (*blue*), while the N-ethyl fragment was engaged in van der Waal's contacts with γP72 (*light green line*) and γD174 residues (Fig. 5C).

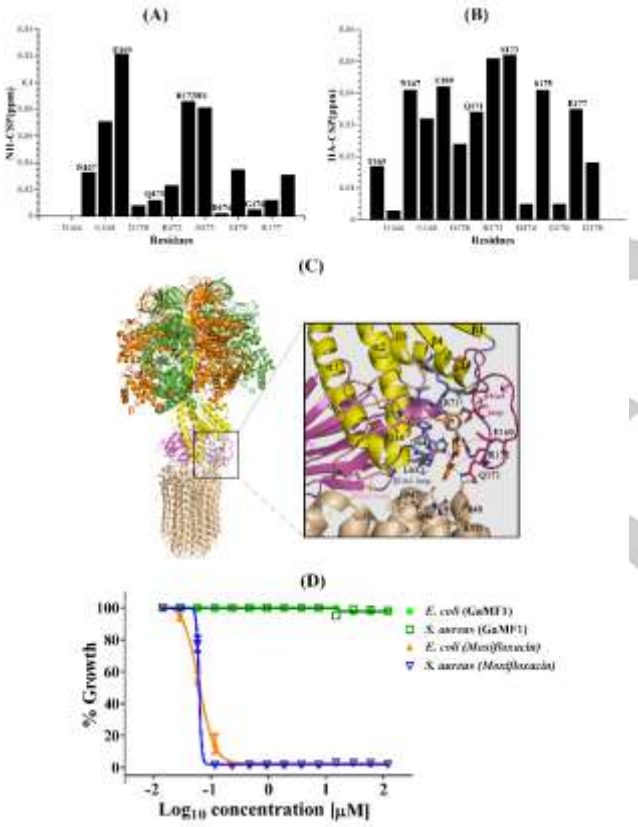


Figure 5: (*A*) 2D-TOCSY NMR spectral data of the free *Mtb* γ166-179 peptide and complexed with GaMF1 highlights the chemical shift changes for NH resonances (absolute values of CSP) between free and complex peptide. (*B*) Absolute values of CSP for HA resonances between free and the GaMF1-bound γ166-179 peptide. (*C*) Predicted binding interaction of GaMF1 on the mycobacterial γ-c ring interface. The three loop regions of β1/α1 (*blue*), β5/α2 (*light red*) and β4/α4 (*magenta*) and together with residues of the *c*-ring loops (*beige*) form the ligand binding pocket. (*D*) A screen of GaMF1 against *S. aureus* and *E. coli*, as representatives of the human microbiome, confirm the specificity of GaMF1 for *Mtb* as well as that the compound does not affect the human microbiome.

To confirm its target specificity and to demonstrate that GaMF1 is not a broad spectrum antibiotic, the compound was tested against *Staphylococcus aureus* and *Escherichia coli* as representatives of Gram-positive and Gram-negative bacteria, respectively (Fig. 5D). While the antibiotic Moxiflocaxin, which is active against both Gram-positive and Gram-negative bacteria by inhibiting DNA gyrase, inhibits both the growth of *S. aureus* and *E. coli*, no growth inhibition of the two bacteria strains was observed in the presence of increasing concentration of GaMF1. Since the extra γ-loop is specific to mycobacterial F-ATP synthases, GaMF1 is not expected to bind to the *E. coli* or *S. aureus* F-ATP synthase and consequently should not inhibit the major energy converter of these prokaryotic cells. NMR titration experiments and docking studies are consistent with a direct interaction between GaMF1 and the subunit *Mt*γ. However, we cannot exclude the possibility that GaMF1 may interfere with additional targets in the pathogen. Future studies could focus on the selection of escape mutants resistant to GaMF1 to confirm target engagement genetically, or alternatively to identify indirect mechanism of resistance such as those described for BDQ of clofazimine [21].

GaMF1 synergizes with BDQ and TBAJ-876 analogs

The *in vitro* ATP synthesis experiments of the I66M mutant IMVs and GaMF1 indicated a synergistic effect but no cross resistance to the diarylquinoline BDQ (Fig. 4F). To assess a synergistic effect of both compounds in more detail, the effect of drug combinations on the growth inhibition of *M. smegmatis* was examined. BDQ inhibited growth at 31.3 nM. However, in combination with GaMF1, the IC_{50} of BDQ reduced with the concentration of GaMF1 added, thus confirming a synergistic effect of both compounds (Fig. 6A). Next, the synergistic effect on ATP synthesis was studied in *M. bovis* BCG. In the whole cell ATP synthesis assay, an additive effect of BDQ and GaMF1 was observed (Fig. 6B). BDQ alone at 200 nM reduced ATP synthesis from $3.7x10^4$ RLU to $8x10^3$ RLU. After addition of GaMF1, ATP formation dropped in a concentration dependent manner of GaMF1 and lowered the IC_{50} of BDQ significantly (Fig. 6B).

Most recently, diarylquinoline analogs with improved potency and physio-chemical properties have been described. [6a-6c] These improved analogues include 5307, 5316, and 5377 [6c] that interact with the subunit *c*-ring and subunit ε of the mycobacterial F-ATP synthase [6d] (Fig. 6C). Using a modified synthesis protocol (see Supplementary *Information*) the racemates of 5307 ((±)-5307), 5316 ((±)-5316), and 5366 ((±)-5366) were synthesized (CCDC deposition numbers: 1967685, 1967683, 1967684, respectively), which inhibited *M. smegmatis* mc^2 155 growth at an MIC_{50} of 1.6 nM, 4.4 nM and 1.0 nM, respectively (Fig. 6D). The improved MIC_{50} values of the racemates compared to BDQ were in line with improved IC_{50} in ATP synthesis inhibition of *M. smegmatis* mc^2 155 IMVs with values of 0.1 ± 0.01 nM, 0.2 ± 0.08 nM, and 0.6 ± 0.33 nM for the racemates 5307, 5316, and 5366, respectively (Fig. 6E). Importantly, combining (±)-5307 with 300 nM of GaMF1 revealed a significant reduction in cell growth (Fig. 6F), demonstrating that the potency of the already improved DARQ derivative increased by the synergistic effect of GaMF1. These results open a new avenue for the development of a novel combinatory drug regimen. Interestingly, the combination of 300 nM GaMF1 with either (±)-5316 or (±)-5366 did not cause a synergistic effect, reflecting a specific combinatorial potency of (±)-5307 and GaMF1.

Viability and pluripotency of human embryonic stem cells remain unaffected by BDQ, (±)-5307 and GaMF1

Human stem cell technologies are becoming increasingly important as valuable *ex vivo* test models during the drug discovery process (reviewed in [22]). We therefore decided to employ a human embryonic stem cell (hESC) line (E3) that we have recently genetically engineered into a highly sensitive pluripotency reporter [23], and examined both potential drug-induced genotoxicity and perturbations of the hESC transcriptional program that lead to hESC differentiation. The assays revealed that neither the tested compounds (±)-5307, (±)-

This article is protected by copyright. All rights reserved.

Accepted Manuscript

5307 in combination with 300 nM GaMF1 nor the control BDQ elicited substantial genotoxic effects or induced major alterations in the global transcriptional program (Supplementary Information; Supplementary Fig. 2A-G).

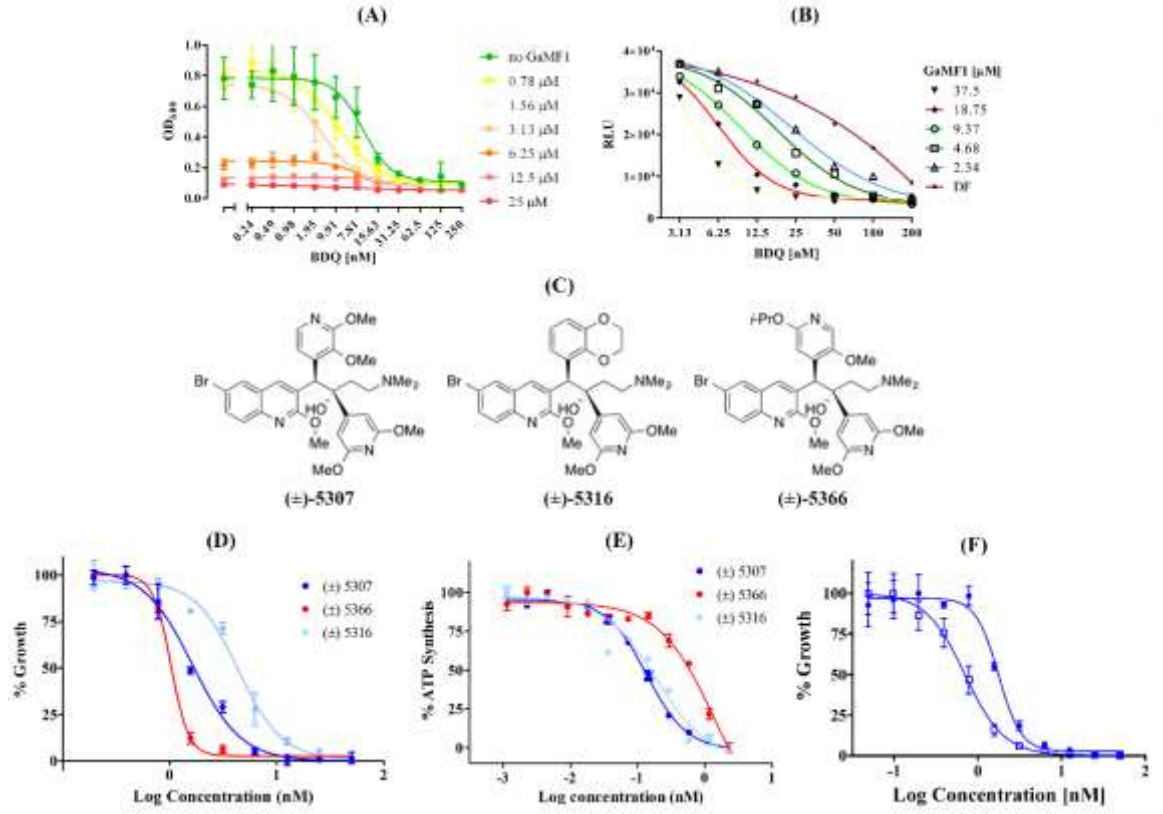


Figure 6: (*A*) Synergistic effect of the TB drug BDQ and GaMF1 on cell growth of *M. smegmatis*. Combination of different concentrations of BDQ and GaMF1 was done by two subsequent serial dilutions in the 96 well plate containing 100 µl of 7H9 media. Suspension of *M. smegmatis* mc^2 155 in logarithmic growth phase was added to obtain starting OD_{600} = 0.05 in 200 µl. Plates were incubated 24 h in 37 °C. (*B*) Whole cell ATP synthesis assay with *M. bovis* BCG in the presence of BDQ and GaMF1. (*C*) Structures of (±)-5307, (±)-5316, and (±)-5366. (*D*) ATP synthesis inhibition of *M. smegmatis* IMVs by (±)-5307, (±)-5316, and (±)-5366, respectively. (*E*) *M. smegmatis* growth inhibition affected by (±)-5307 (●) and in combination with 300 nM of GaMF1 (□).

SAR studies of GaMF1 series

In order to improve the efficacy of GaMF1, a series of analogs was synthesized (see Supplementary Information) and tested in mycobacterial ATP synthesis assay (IC_{50}) and in the mycobacterial assay (*M. smegmatis* and/or *Mtb* MIC_{50}). Key SAR findings are summarized in Table 1.

It was established that ring C was essential for maintaining activity as truncated analog **1** (Table 1) did not inhibit ATP synthesis (IC_{50} > 100 µM). The amide functionality linking rings A and B was essential for activity as switching to the reverse amide **2** led to a loss in activity. This is consistent with the docking pose presented previously (Fig. 5C), where the carbonyl and amide atoms on ring A were predicted to be in close contacts with γQ171 and γD64.

While the removal of substitution on the phenyl ring (**3**) led to a modest drop in activity, replacement of the 3-bromo with either electron-withdrawing CF_3 (**4**) or electron-donating pyrrolidine (**5**) substituents gave compounds with comparable potencies. This result suggests that lipophilic substitutions are preferred in this part of the molecule and that the electronics of the aromatic ring have little contribution towards potency. This observation is also consistent with the docking pose (Fig. 5C), which shows (i) space to tolerate substituents on the phenyl ring and (ii) no key interactions between the neighbouring residues and the phenyl ring. Hence, we concluded that the region around phenyl ring A would perhaps be the most amendable to structural modifications.

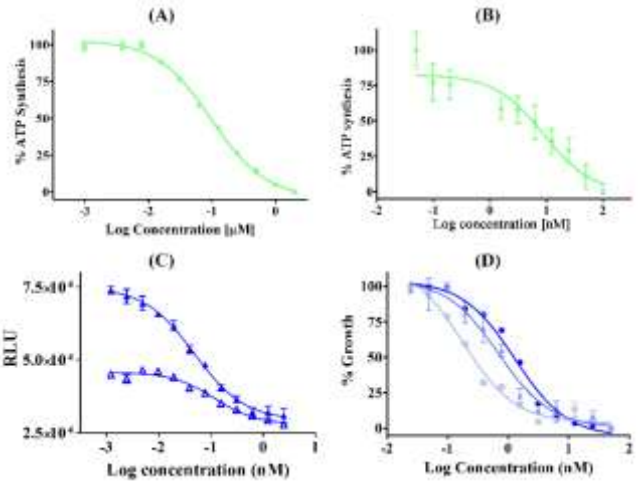


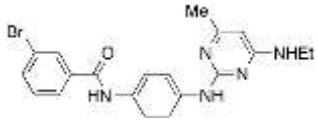
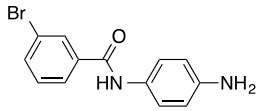
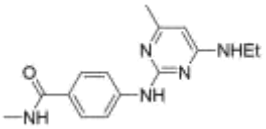
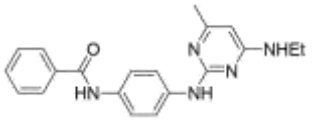
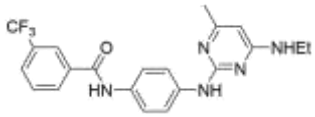
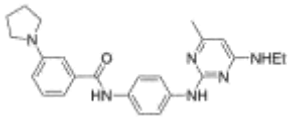
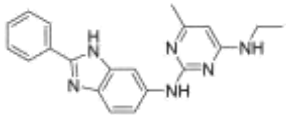
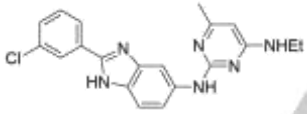
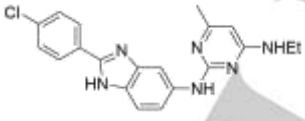
Figure 7: Inhibition of ATP synthesis by GaMF1 analog **8** in IMVs from *M. smegmatis* with an IC_{50} of 90 nM on *M. smegmatis* (*A*) and 8.7 nM on *M. bovis* BCG IMVs (*B*). (*C*) 300 nM of GaMF1 analog **8** increase the potency of (±)-5307 in ATP synthesis inhibition (Δ) when compared to the inhibitory effect of (±)-5307 alone (▲). (*D*) Addition of 500 nM (▲) and 9 µM (■) of GaMF1 analog **8** shifted the inhibition curve to the left, showing a 2-fold (0.7 nM) and 10-fold (0.17 nM) decrease in MIC_{50} thereby revealing the synergistic effect on the same target enzyme. The effect of racemate ((±)-5307) on growth of *M. smegmatis* mc^2 155 is shown in curve (●).

This article is protected by copyright. All rights reserved.

Accepted Manuscript

10.1002/anie.202002546

Table 1: SAR of selected GaMF1 analogs

Cpd No.	Structure	IC_{50} (μm)[a]	MIC_{50} (μM)[b]	MIC_{50} (μM)[c]
GaMF1		0.5	11	33
1		>100	>100	-
2		>100	-	-
3		1.7	27	-
4		0.23	13.4	10.1
5		0.21	9.7	13.9
6		0.28	23.5	8.8
7		0.29	14.6	5.2
8		0.09	9.9	3

[a] ATP synthesis Inhibitory Concentration on *M. smegmatis* IMVs
[b] *M. smegmatis* Minimum Inhibitory Concentration
[c] *M. tuberculosis* Minimum Inhibitory Concentration

In an attempt to retain the key amide functionality while addressing potential concerns about toxicity arising due to oxidation of the 1,4-phenylenediamine moiety (ring B), analogs (**6** – **8**) bearing a central benzimidazole motif with a phenyl substituent were synthesized. Such a modification would restrict rotational freedom between the A and B ring while retaining the phenyl substituent. All three analogs displayed good IC_{50} and MIC_{50} values, with one (**8**) showing a 10-fold improvement compared to the original compound, GaMF1. The *p*-chloro phenyl derivative **8** afforded a 18-fold enhancement in ATP synthesis inhibition of about 90 nM on *M. smegmatis* (Fig. 7A) and even 8.7 nM on *M. bovis* BCG IMVs (Fig. 7B) and inhibited *Mtb* growth culture at a concentration of 3 μM, while the *m*-chloro phenyl derivative **7** maintained an activity profile similar to the parent molecule. Interestingly, when 300 nM of analog **8** was combined in an ATP synthesis assay of *M. smegmatis* with (±)-5307 (Fig. 7C) a clear synergistic effect could be observed, reflecting the potency of this unique TB-compound combination. Addition of GaMF1 analog **8** shifted the growth curve to the left, thereby revealing the synergistic effect on the same target enzyme (Fig. 7D).

Conclusion

Respiration is a promising target for the development of new antimycobacterial agents, including the inhibitors clofazimine, Telacebec (Q203), and BDQ which inhibit the mycobacterial complex II (Ndh2; [24]), the cytochrome *bc1* complex [25], and the F-ATP synthase.[3] The latter was approved by FDA for the treatment of MDR-TB in 2012. [4a] An increased understanding of the function and structure of different respiratory complexes and F-ATP synthase in mycobacterial biology is becoming key to advancing cellular bioenergetics as a new target space. [11-13, 14a, 26] Here, we identified a novel TB-compound GaMF1 targeting the unique mycobacterial F-ATP synthase γ-loop and being specific. It retained ATP synthesis inhibition against a BDQ resistant mutant and killing potency against H- and R monoresistant *Mtb* strains as well as the HR-resistant *Mtb* strains. NMR- and docking results of GaMF1 support the mode of action, describing that binding of the compound to the mycobacterial γ-loop interferes with the smooth rotation of the γ subunit along the loop residues of the rotating *c*-ring turbine, thereby interrupting proper coupling of proton conduction in the *a*-*c* subunits and ATP synthesis in the α_3:β_3 headpiece via the central stalk subunit γ. In comparison, BDQ, TBAJ-876 and its derivatives 5307, 5316 and 5377 bind to both subunit *c* [14b] and ε [6d, 14d, 27], and drug binding presumably results in stalling of *c*-ring rotation relative to subunit *a*. [14b] Combining GaMF1 with BDQ or (±)-5307 increased the MIC- and IC potency, while not affecting viability and pluripotency of human embryonic stem cells, which is essential for further clinical applications. Medicinal chemistry efforts resulted in GaMF1 derivatives with increased potency regarding ATP synthesis inhibition and *Mtb* cell viability. Future studies will focus on the utility of the new lead for the treatment of TB, considering that in patients, *Mtb* grows extra- as well as intra-cellular. Here, we demonstrated activity of GaMF1 against *Mtb* growing in broth, i.e. against the extracellular form of the bacterium. Therefore, activity of the new F-ATP synthase inhibitor against intracellular *Mtb* will be determined in macrophage infection models.

Acknowledgements

This research was supported by the National Research Foundation (NRF) Singapore, NRF Competitive Research Programme (CRP), Grant Award Number NRF–CRP18–2017–01. A. Hotra is grateful to receive an IGS Premium Scholarship, Institute of Technology in Health and Medicine at NTU, and J. P. Sarathy received a PhD scholarship from the School of Medicine, NUS, Singapore. We thank Dr. Yongxin Li for the X-ray crystallographic structure determinations.

Conflict of Interest

G.G., R.W.B., A.H., T.D., K.P. are inventors on the patent PCT/SG2018/050075, and G.G., R.W.B., P.S.N. are inventors on the patent 10201911205R which are related to the chemical series described in this article.

This article is protected by copyright. All rights reserved.

Accepted Manuscript

Keywords: ATP synthesis • Bioenergetics • Drug Discovery • F-ATP synthase • OXPHOS • Tuberculosis

References

[1] World Health Organization (2019) Global tuberculosis report 2019. *https://www.who.int/tb/publications/global_report/en/*.

[2] a) G. M. Cook, K. Hards, C. Vilcheze, T. Hartman, M. Berney, *Microbiol Spectr* **2014**, *2* doi: 10.1128/microbiolspec; b) S. P. Rao, S. Alonso, L. Rand, T. Dick, K. Pethe, *Proc Natl Acad Sci U S A* **2008**, *105*, 11945-11950.

[3] K. Andries, P. Verhasselt, J. Guillemont, H. W. Gohlmann, J. M. Neefs, H. Winkler, J. Van Gestel, P. Timmerman, M. Zhu, E. Lee, P. Williams, D. de Chaffoy, E. Huitric, S. Hoffner, E. Cambau, C. Truffot-Pernot, N. Lounis, V. Jarlier, *Science* **2005**, *307*, 223-227.

[4] a) K. Andries, C. Villellas, N. Coeck, K. Thys, T. Gevers, L. Vranckx, N. Lounis, B. C. de Jong, A. Koul, *PLoS One* **2014**, *9*, e102135; b) R. C. Hartkoorn, S. Uplekar, S. T. Cole, *Antimicrob Agents Chemother* **2014**, *58*, 2979-2981.

[5] a) A. H. Diacon, P. R. Donald, A. Pym, M. Grobusch, R. F. Patientia, R. Mahanyele, N. Bantubani, R. Narasimooloo, T. De Marez, R. van Heeswijk, N. Lounis, P. Meyvisch, K. Andries, D. F. McNeeley, *Antimicrob Agents Chemother* **2012**, *56*, 3271-3276; b) J. Guillemont, C. Meyer, A. Poncelet, X. Bourdrez, K. Andries, *Future medicinal chemistry* **2011**, *3*, 1345-1360; cFDA, **2012**.

[6] a) A. S. T. Tong, P. J. Choi, A. Blaser, H. S. Sutherland, S. K. Y. Tsang, J. Guillemont, M. Motte, C. B. Cooper, K. Andries, W. Van den Broeck, S. G. Franzblau, A. M. Upton, W. A. Denny, B. D. Palmer, D. Conole, *ACS Med Chem Lett* **2017**, *8*, 1019-1024; b) P. J. Choi, H. S. Sutherland, A. S. T. Tong, A. Blaser, S. G. Franzblau, C. B. Cooper, M. U. Lotlikar, A. M. Upton, J. Guillemont, M. Motte, L. Queguiner, K. Andries, W. Van den Broeck, W. A. Denny, B. D. Palmer, *Bioorg Med Chem Lett* **2017**, *27*, 5190-5196; c) H. S. Sutherland, A. S. T. Tong, P. J. Choi, D. Conole, A. Blaser, S. G. Franzblau, C. B. Cooper, A. M. Upton, M. U. Lotlikar, W. A. Denny, B. D. Palmer, *Bioorganic & medicinal chemistry* **2018**, *26*, 1797-1809; d) J. P. Sarathy, P. Ragunathan, J. Shin, C. B. Cooper, A. M. Upton, G. Grüber, T. Dick, *Antimicrob Agents Chemother* **2019**, *63* pii: e01191-19.

[7] S. J. Tantry, S. D. Markad, V. Shinde, J. Bhat, G. Balakrishnan, A. K. Gupta, A. Ambady, A. Raichurkar, C. Kedari, S. Sharma, N. V. Mudugal, A. Narayan, C. N. Naveen Kumar, R. Nanduri, S. Bharath, J. Reddy, V. Panduga, K. R. Prabhakar, K. Kandaswamy, R. Saralaya, P. Kaur, N. Dinesh, S. Guptha, K. Rich, D. Murray, H. Plant, M. Preston, H. Ashton, D. Plant, J. Walsh, P. Alcock, K. Naylor, M. Collier, J. Whiteaker, R. E. McLaughlin, M. Mallya, M. Panda, S. Rudrapatna, V. Ramachandran, R. Shandil, V. K. Sambandamurthy, K. Mdluli, C. B. Cooper, H. Rubin, T. Yano, P. Iyer, S. Narayanan, S. Kavanagh, K. Mukherjee, V. Balasubramanian, V. P. Hosagrahara, S. Solapure, S. Ravishankar, P. S. Hameed, *J Med Chem* **2017**, *60*, 1379-1399.

[8] S. Kumar, R. Mehra, S. Sharma, N. P. Bokolia, D. Raina, A. Nargotra, P. P. Singh, I. A. Khan, *Tuberculosis* **2018**, *108*, 56-63.

[9] W. G. Saw, M. L. Wu, P. Ragunathan, G. Biuković, A. M. Lau, J. Shin, A. Harikishore, C. Y. Cheung, K. Hards, J. P. Sarathy, R. W. Bates, G. M. Cook, T. Dick, G. Grüber, *Scientific Rep.* **2019**, *9*, 16759.

[10] a) A. Mandavilli, *Nature medicine* **2007**, *13*, 271; b) D. G. Russell, C. E. Barry, 3rd, J. L. Flynn, *Science* **2010**, *328*, 852-856.

[11] A. Koul, E. Arnoult, N. Lounis, J. Guillemont, K. Andries, *Nature* **2011**, *469*, 483-490.

[12] C. M. Sassetti, D. H. Boyd, E. J. Rubin, *Mol Microbiol* **2003**, *48*, 77-84.

[13] G. M. Cook, K. Hards, E. Dunn, A. Heikal, Y. Nakatani, C. Greening, D. C. Crick, F. L. Fontes, K. Pethe, E. Hasenoehrl, M. Berney, *Microbiol Spectr* **2017**, *5*.

[14] a) A. T. Zhang, M. G. Montgomery, A. G. W. Leslie, G. M. Cook, J. E. Walker, *Proc Natl Acad Sci U S A* **2019**; b) L. Preiss, J. D. Langer, O. Yildiz, L. Eckhardt-Strelau, J. E. Guillemont, A. Koul, T. Meier, *Sci Adv* **2015**, *1*, e1500106; c) N. Kamariah, R. G. Huber, W. Nartey, S. Bhushan, P. J. Bond, G. Grüber, *J Struct Biol* **2019**, *207*, 199-208; d) S. Joon, P. Ragunathan, L. Sundararaman, W. Nartey, S. Kundu, M. S. S. Manimekalai, N. Bogdanović, T. Dick, G. Grüber, *FEBS J* **2018**, *285*, 1111-1128; e) N. Kamariah, P. Ragunathan, J. Shin, W. G. Saw, C. F. Wong, T. Dick, G. Grüber, *Progr. Biophysi. Mol. Biol.* **2019**, 152, 64-73.

[15] A. Hotra, M. Suter, G. Biuković, P. Ragunathan, S. Kundu, T. Dick, G. Grüber, *FEBS J* **2016**, *283*, 1947-1961.

[16] R. Priya, G. Biuković, M. S. Manimekalai, J. Lim, S. P. Rao, G. Grüber, *J Bioenerg Biomembr* **2013**, *45*, 121-129.

[17] A. Koul, L. Vranckx, N. Dhar, H. W. Gohlmann, E. Ozdemir, J. M. Neefs, M. Schulz, P. Lu, E. Mortz, J. D. McKinney, K. Andries, D. Bald, *Nat Commun* **2014**, *5*, 3369.

[18] P. Ragunathan, H. Sielaff, L. Sundararaman, G. Biuković, M. S. Subramanian Manimekalai, D. Singh, S. Kundu, T. Wohland, W. Frasch, T. Dick, G. Grüber, *J Biol Chem* **2017**, *292*, 11262-11279.

[19] N. P. Kalia, E. J. Hasenoehrl, N. B. Ab Rahman, V. H. Koh, M. L. T. Ang, D. R. Sajorda, K. Hards, G. Grüber, S. Alonso, G. M. Cook, M. Berney, K. Pethe, *Proc Natl Acad Sci U S A* **2017**, *114*, 7426-7431.

[20] a) S. Petrella, E. Cambau, A. Chauffour, K. Andries, V. Jarlier, W. Sougakoff, *Antimicrob Agents Chemother* **2006**, *50*, 2853-2856; b) S. Kundu, G. Biuković, G. Grüber, T. Dick, *Antimicrob Agents Chemother* **2016**, *60*, 6977-6979.

[21] R. C. Hartkoorn, S. Uplekar, S. T. Cole, Antimicrob Agents Chemother **2014**, *58*, 2979-2981.

[22] a) K. G. Chen, B. S. Mallon, K. Park, P. G. Robey, R. D. G. McKay, M. M. Gottesman, W. Zheng, *Trends in molecular medicine* **2018**, *24*, 805-820; b) Y. Avior, I. Sagi, N. Benvenisty, *Nature reviews. Molecular cell biology* **2016**, *17*, 170-182.

[23] S. H. Vijaya Chandra, H. Makhija, S. Peter, C. M. Myint Wai, J. Li, J. Zhu, Z. Ren, M. S. D'Alcontres, J. W. Siau, S. Chee, F. J. Ghadessy, P. Dröge, *Nucleic Acids Res* **2016**, *44*, e55.

[24] T. Yano, S. Kassovska-Bratinova, J. S. Teh, J. Winkler, K. Sullivan, A. Isaacs, N. M. Schechter, H. Rubin, *J Biol Chem* **2011**, *286*, 10276-10287.

[25] K. Pethe, P. Bifani, J. Jang, S. Kang, S. Park, S. Ahn, J. Jiricek, J. Jung, H. K. Jeon, J. Cechetto, T. Christophe, H. Lee, M. Kempf, M. Jackson, A. J. Lenaerts, H. Pham, V. Jones, M. J. Seo, Y. M. Kim, M. Seo, J. J. Seo, D. Park, Y. Ko, I. Choi, R. Kim, S. Y. Kim, S. Lim, S. A. Yim, J. Nam, H. Kang, H. Kwon, C. T. Oh, Y. Cho, Y. Jang, J. Kim, A. Chua, B. H. Tan, M. B. Nanjundappa, S. P. Rao, W. S. Barnes, R. Wintjens, J. R. Walker, S. Alonso, S. Lee, J. Kim, S. Oh, T. Oh, U. Nehrbass, S. J. Han, Z. No, J. Lee, P. Brodin, S. N. Cho, K. Nam, J. Kim, *Nature medicine* **2013**, *19*, 1157-1160.

[26] a) B. Wiseman, R. G. Nitharwal, O. Fedotovskaya, J. Schafer, H. Guo, Q. Kuang, S. Benlekbir, D. Sjostrand, P. Adelroth, J. L. Rubinstein, P. Brzezinski, M. Hogbom, *Nat Struct Mol Biol* **2018**, *25*, 1128-1136; b) H. Gong, J. Li, A. Xu, Y. Tang, W. Ji, R. Gao, S. Wang, L. Yu, C. Tian, J. Li, H. Y. Yen, S. Man Lam, G. Shui, X. Yang, Y. Sun, X. Li, M. Jia, C. Yang, B. Jiang, Z. Lou, C. V. Robinson, L. L. Wong, L. W. Guddat, F. Sun, Q. Wang, Z. Rao, *Science* **2018**, *362*; c) W. Moreira, D. B. Aziz, T. Dick, *Frontiers in microbiology* **2016**, *7*, 199.

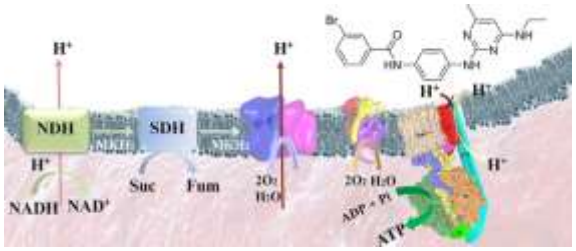
[27] G. Biuković, S. Basak, M. S. Manimekalai, S. Rishikesan, M. Roessle, T. Dick, S. P. Rao, C. Hunke, G. Grüber, *Antimicrob Agents Chemother* **2013**, *57*, 168-176.

Accepted Manuscript

This article is protected by copyright. All rights reserved.

10.1002/anie.202002546

Entry for the Table of Contents



The antimycobacterial GaMF1 inhibits the mycobacterial F-ATP synthase by binding to a specific loop of subunit γ, preventing ATP synthesis via oxidative phosphorylation. The bactericidal compound is active against multidrug-resistant strains, increased the potency of the TB-drug Bedaquiline and other DARQ analogues, opening a new avenue for an efficient multi-drug combination. Chemistry efforts resulted in analogues with nanomolar potencies.

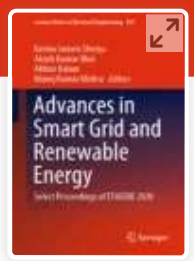
WILEY-VCH

Accepted Manuscript

This article is protected by copyright. All rights reserved.

Springer Link

Search Log in



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 539–546

Digital Art Using Machine Learning

Karri Yaswanth Teja Reddy, Karnati Madhava Sai Kumar, A. Pravin, T. Prem Jacob & G. Nagarajan

Conference paper | First Online: 05 January 2021

309 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Digital art creation uses digital art that includes multiple colors support for several applications. Digital art implies making out of semantic characters in a could not capacity by hand or finger advancements. Digital arts contrast from customary scribble, the last consists of object-up-object-down development, when the past needs a determined plan of creating occurrence. We label digital art affirmation issues in a few accomplice papers. For some type of exceptions, the data which have been drawn through hand signals were processed or trained in a easy way. We label digital art on two

levels: development characters and multicolored characters. Detached digital art characters could be seen like development movements regardless of the way that with extended intricacy and variance. In order to avoid overlapping of letters, we created different types of characters and stored in a virtual space for exhibiting the data. A covered Markov model is employed for digital art exhibiting and affirmation. We display the event data along estimations past a two-dimensional bearing might be helpfully particular for digital art affirmation. We display the development data along estimations past a two-dimensional bearing could be helpfully particular for digital art affirmation. Now, look into the general ampleness of various component estimations of optical and inertial after signs and resulting the plausible affirmation execution accordingly. The proposed structure attains a characters screw up pace of 0.8% for characters-based affirmation and 1.9% for different color-based affirmation.

Keywords

Digital art **Intra class variability**

Machine learning **Image processing**

Markov model

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_53
- Chapter length: 8 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Price includes VAT (India)

- ISBN: 978-981-15-7511-2
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7714-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7510-5
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. B. Bharathi, P. Vijayalakshmi, T. Nagarajan, Speaker identification using utterances correspond to speaker-specific-text, in *IEEE Technology Students' Symposium*. IEEE, Jan 2011, pp. 171–174

2. J., Sujolincy, M.D. Kamalesh, An authenticated smart card service using cross-cloud and big data, in *2016 International Conference on Control, Instrumentation, Communication and Computational Technologies* (*ICCICCT*). IEEE, Dec 2016, pp. 634–638

3. P.V. Raja, K. Sangeetha, D. Deepa, Extractive text summarization system using fuzzy clustering algorithm for mobile devices. Asian J. Inf. Technol. **15**(5), 933–939 (2016)

4. M. Chen, G. AlRegib, B.-H. Juang, Air-making affirmation—part II: identification and affirmation of making activity in consistent stream of development data. IEEE Trans. Hum. Mach. Syst. to be disseminated

5. S.P. Shyry, Novel enhanced encryption algorithm for shared key generation, in *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing*, Oct 2014, pp. 1–7

6. M. Chen, G. AlRegib, B.-H. Juang, Feature planning and showing for 6d development movement affirmation. IEEE Trans. Blend. Med. **15**(3), 561–571 (2013)

7. C.H. Blickenstorfer, Graffiti: Amazing! Pen Process. Mag. **1**, 30–31 (1995)

8. K. Perlin, Quikwriting: constant stylus-based substance section, in *Proceedings of the Eleventh Annual ACM Symposium on UI Software and Technology* (1998), pp. 215–216.

9. R.S.B. Krishna, M. Aramudhan, Unsupervised spectral sparse regression feature selection using social media datasets, in *Proceedings of the International Conference on Informatics and Analytics*, Aug 2016, pp. 1–5

10. P. Isokoski, R. Raisamo, Quikwriting as a multi-contraption content area procedure, in *Proceedings of the Third Nordic Conference on Human-PC Connection* (2004), pp. 105–108

11. T. Költringer, P. Isokoski, T. Grechenig, Twostick: composing with a game controller. Proc. Graph. Interface, 103–110 (2007)

12. N. Srinivasan, C. Lakshmi, A novel prediction based tree structured data using machine learning techniques. Res. J. Pharm. Biol. Chem. Sci. **7**(5), 527–531 (2016)

13. R. Plamondon, S. Srihari, On the web and disengaged scribble affirmation: an extensive diagram. IEEE Trans. Model Butt-Driven. Mach. Intell. **22**(1), 63–84 (2000)

14. J. Makhoul, T. Starner, R. Schwartz, G. Chou, On-line cursive hand-making affirmation using hid markov models and quantifiable gram-mutilates. Proc

15. Hop development (2015) [Online]. Available: https://www.leapmotion.com

16. Swype—type brisk, swype speedier (2013) [Online]. Available: https://www.swype.com/

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Karri Yaswanth Teja Reddy, Karnati Madhava Sai Kumar, A. Pravin, T. Prem Jacob & G. Nagarajan

Corresponding author

Correspondence to Karri Yaswanth Teja Reddy .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Yaswanth Teja Reddy, K., Sai Kumar, K.M., Pravin, A., Prem Jacob, T., Nagarajan, G. (2021). Digital Art Using Machine Learning. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_53

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7511-2_53

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)
IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0

Diagnosis of Liver Disease using Machine Learning Models

A.Sivasangari
School of Computing
Sathyabama Institute of Science and Technology
Chennai, India
sivasangarikavya@gmail.com

Baddigam Jaya Krishna Reddy
School of Computing
Sathyabama Institute of Science and *Technology*
Chennai, India
jayakrishnareddy728@gmail.com

Annamareddy Kiran
School of Computing
Sathyabama Institute of Science and Technology
Chennai, India
kiranchowdary1601@gmail.com

P.Ajitha
School of Computing
Sathyabama Institute of Science and Technology
Chennai, India
hannahgracelyne@gmail.comail.com

***Abstract*— Liver disease is one of the key causes of high numbers of deaths in the country and is considered a life-threatening disease, not just anywhere, but worldwide. Liver disease can also impact peoples early in their life. More than 2.4 per cent of annual Indian deaths are due to liver disorders. It is also difficult to detect liver disease due to mild symptoms in the early stages. If it is too late the signs always come to light. Thus liver-related disease poses more problems for people living and is more important nowadays to recognize the causes, and identification phase. So, for early detection of liver disease, an automated program is needed to build with more accuracy and reliability. Specific machine learning models are developed for this purpose to predict the disease. In this paper, the methods of Support Vector Machines (SVM), Decision Tree (DT) and Random Forest (RF) is proposed to predict liver disease with better precision, accuracy and reliability.**

Keywords: Liver Infection, Random Forest, Confusion Matrix.

I. INTRODUCTION

Liver infection is a precarious malady to analyze given the nuance of the indications while in the beginning periods. Issues with liver illnesses are not found until it is regularly past the point of no return as the liver keeps on working in any event, when incompletely harmed. Early determination can conceivably be life-sparing. While it is difficult to diagnose even the experienced herbal practitioner, the early signs of these diseases that be identified. Late patient experiences will graciously build up his / her standards of life. Therefore the findings of this analysis are important from the point of view of both the computer scientist and the medical professional.

This paper proposes the three machine learning models, i.e. help for Vector Machines, Decision Tree and Random Forest for the prediction of liver disease.

II RELATED WORKS

In [1], Vasan Durai et al presented a liver disease prediction based on the machine learning model. The proposed supervised classification algorithm J48 algorithm for prediction. The dataset is collected from the UCI repository dataset. Nazmun et al proposed liver disease classification by using decision tree classifier. It achieved higher accuracy than other algorithms[2].

In [3], Ramalingam et al proposed the machine learning model for prediction of liver-related diseases. For this work, liver data was collected from UCI which is related to hepatitis and hepatocellular carcinoma liver disease. They proposed the K means algorithm, SVM, LR, RF and Neural Networks for liver disease prediction.

Joel Jacob et al presented a novel approach for liver disease detection. The Indian Liver Patient Dataset (ILPD), picked from the learning repository of UCI, was retrieved for this analysis. They implemented the Regression model, KNN model and ANN model liver disease detection.

Shambel et al in [5] proposed a data mining technique to predict and analyze the liver disorder. They used SVM, decision tree classifiers. They performed data partitioning based on test set to test the model. Dataset repository UCI is used for prediction. Comparison of the type of liver disease discussed at [6] and [7]. In[8] Ignisha Rajathi et al proposed the hybrid WOA-SA and ensemble classifier was proposed. The performance analysis can be done by using CT image dataset. They proposed the method got better accuracy, sensitivity and specificity. [9],[10],[11],[12] Different IoT devices are used in different applications such as healthcare, smart city.

III PROPOSED WORK

This paper presents the three distinct machine algorithms i.e. SVM, DT, and RF to predict or classify the liver disease by using different attributes. Fig.1 describes the workflow of the proposed method for liver disease prediction, as follows,

1. UCI machine learning library gathers data from the liver disease dataset.
2. The dataset has some -1 values; numerical values are required to replace it. This process is carried out in the pre-processing phase.
3. Our data set is divided into data for validation for the training and testing process.
4. Finally trained data can be evaluated with various algorithms and test data are graded based on a trained model.

1. Dataset Collection

UCI machine learning library includes the Patient Dataset for Indian Liver. The data set includes 416 reports of liver disease patients and 167 records of nonliver disease patients. The data collection was collected from Andhra Pradesh in north-eastern India. The data collection contains 441 records for males patients and 142 records for females patients. 10 attributes are being used such as age of the person, gender, total bilirubin, direct billuubin, direct bilirubin, alkaline phosphotase, alamine, aspartate aminotransferase, total protein , albumin and albumin ratio, and the liver disease prediction globulin ratio.

2. Data Preprocessing

The dataset is composed of original attributes and -1 values. The -1 values cannot be processed in programming This transforms these values into a different value, i.e. a numerical value. -1 Column mean values are replaced.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:36:32 UTC from IEEE Xplore. Restrictions apply.

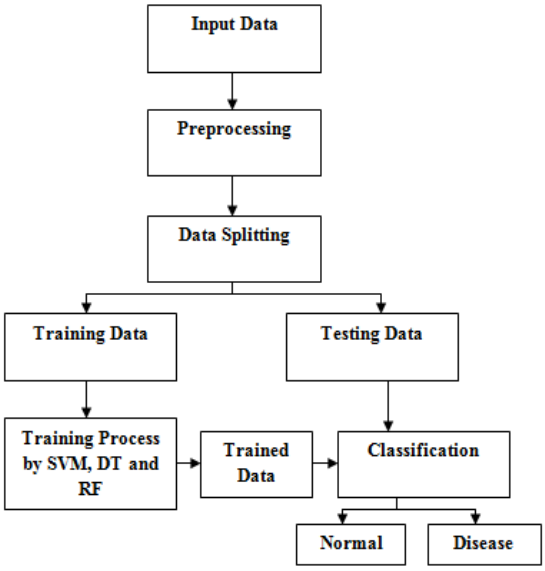


Fig.1 Overall Process for Liver Disease Prediction

3. Data Splitting

Throughout data slicing, the entire dataset is divided into data for the training and testing of the process analysis. In that, 80 percent of data is used for demo data for training and 20 percent for testing.

4. Classification

In classification, accurate detection of disease by using training and testing data set. It proposes three models of machine learning to build the prediction. First training data is trained across three models of machine learning i.e. Decision Tree, Random Forest and Support Vector Machine are predicted on the basis of a trained model of learning, one by one, and then test data. Some parameters including accuracy, precision, and recall are finally compared with the above algorithms. The details given below are three different algorithms,

A. Support Vector Machine

SVM is one of the most widely used computer supervised learning model to perform prediction and classification. SVM has used the hyperplane in the function space that differentiates the labels or groups. An SVM model interprets the training data points as points within the function domain, distributed in a way that distinguishes as broadly as possible the points belonging to different classes the same area, the test data points are then drawn and graded by which side of the threshold they fall.

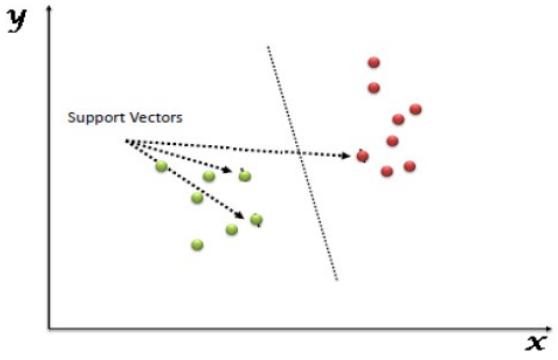


Fig.2 Support Vector Machine

B. Decision Tree

Decision Tree is one of the Supervised Learning algorithms. Classification issues are mostly dealt with by the use of a decision tree. It works easily with attributes which are constant and categorical. Based on important predictors, the population divides into two or more related DT sets. The DT's first step is to calculate entropy for each and every attribute. Next, the dataset is split with high information gain or less entropy based on the variables/predictors. Remaining attributes are followed by two steps above.

$$Entropy\,(E) = \sum_{k=1}^{l} -q_k \log_2 q_k \qquad (1)$$

Where l the variable response modules count is referred to, q_k is the ratio of the number of kth class methods to the total number of models.

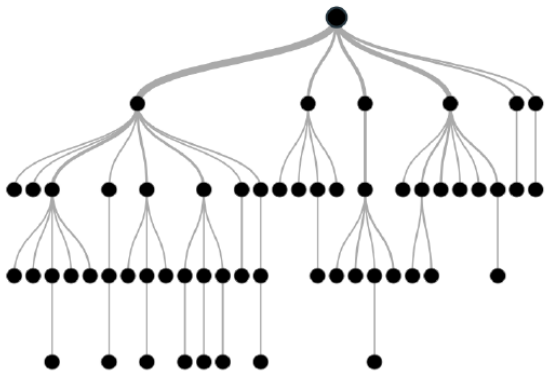


Fig.3. Decision Tree Structure

C. Random Forest

Random Forest is likewise a typical approach of supervised machine learning applications. RF is conducted both for the task of classification but only gives a better classification result. Therefore, the random forest refers to the combination of various trees for decision. Large numbers of trees will do a positive result in RF. The voting method is used to define and then decide on the class, thus allowing for mean estimation in regression for all results of each and every decision trees. Random forest operated very fast and efficiently, with more high dimensionality data set.

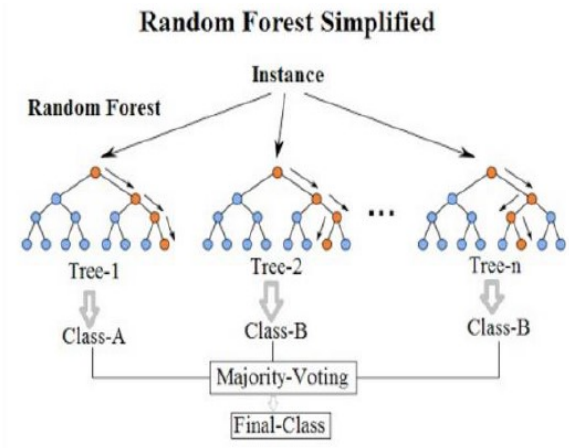


Fig.4. Random Forest Working Structure

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:36:32 UTC from IEEE Xplore. Restrictions apply.

IV EXPERIMENTAL RESULTS

In this section, the categorized outcome from various prediction models was being shown. For contrast with different ones, different parameters; i.e. the parameters. Precise, Precise, and Recall is used.

Indian Liver Patient Dataset contains the total data of 583 patients. 500 patient data is used to train in that, and 83 patient data to test.

Table 1: Confusion Matrix

Actual/Predicted	Normal	Abnormal
Normal	TP_{value}	FN_{value}
Abnormal	FP_{value}	TN_{value}

Table 2: Confusion Matrix of SVM

Actual/Predicted	Normal	Disease	Total
Normal	17	0	17
Disease	4	62	66

Table 3: Confusion Matrix of DT

Actual/Predicted	Normal	Disease	Total
Normal	11	6	17
Disease	4	62	66

Table 4: Confusion Matrix of RF

Actual/Predicted	Normal	Disease	Total
Normal	11	6	17
Disease	0	66	66

Based on the above prediction, the parameters i.e. accuracy, precision and recall are evaluated.

i) Accuracy: This compares the TP and TN analytics to the total number of test images.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

ii) Precision: It is the true positive estimation measurement to the aggregate value of the true positive and false positive values. It is represented in Eqn. (3)

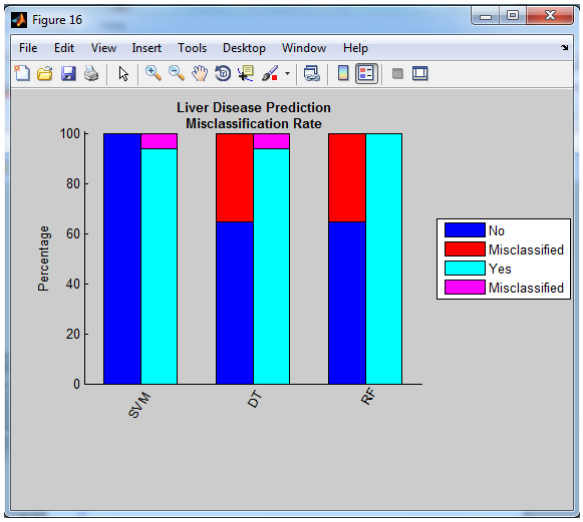
$$Precision = \frac{(TP)}{(TP+FP)} \quad (3)$$

iii) Recall: It is the estimation of the true positive rate to the aggregate value of the true negative rate, both positive and false. It is represented in Eqn. (4).

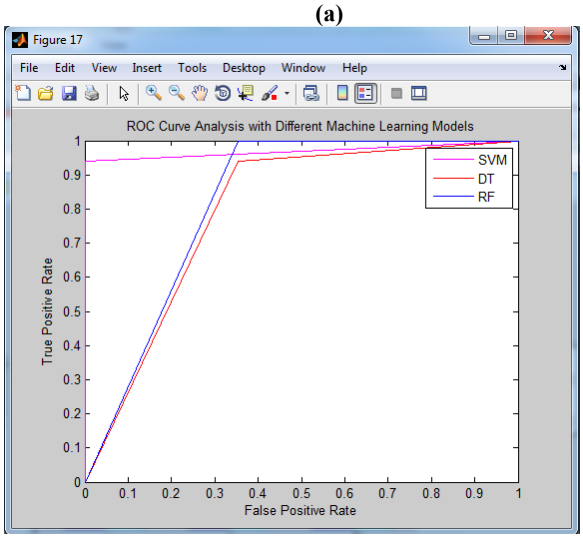
$$Recall = \frac{(TP)}{(TP+FN)} \quad (4)$$

Table 5: Quantitative Evaluation with different models

Method	SVM	DT	RF
Accuracy	0.9518	0.8795	0.9277
Precision	0.8095	0.7333	1.0000
Recall	1.0000	0.6470	0.6470

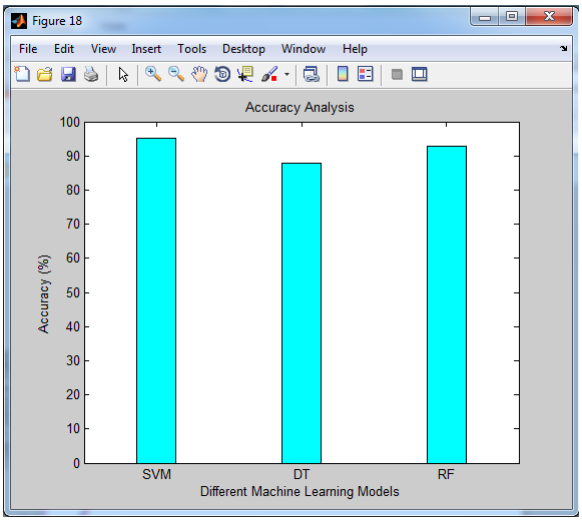


(a)

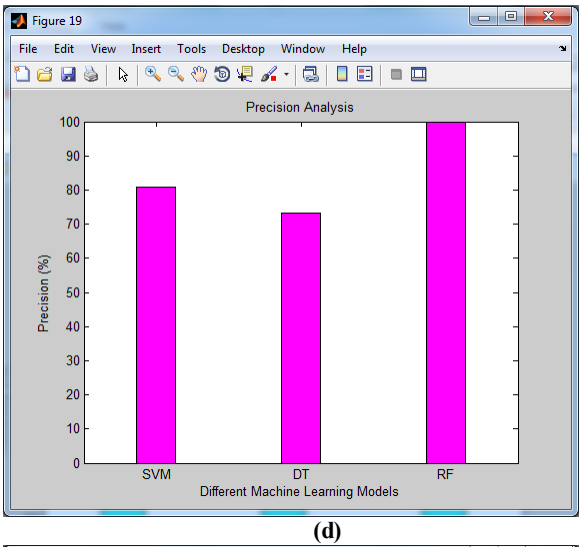


(b)

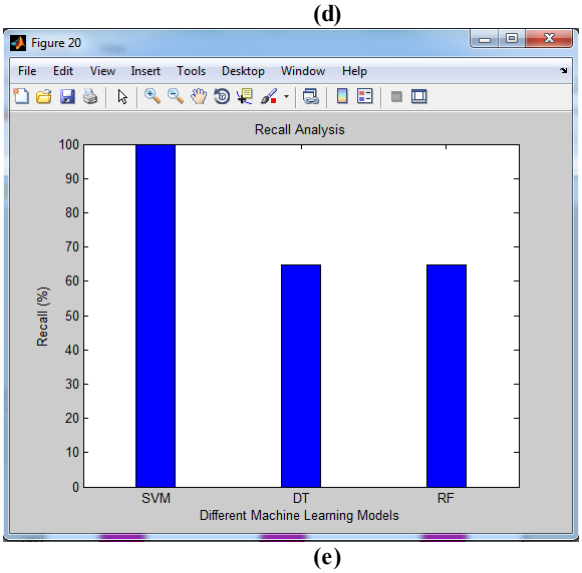
Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:36:32 UTC from IEEE Xplore. Restrictions apply.



(c)



(d)



(e)

Fig.5.(a) Misclassification Rate Prediction. (b) ROC Curve with Different Models. (c) Accuracy Analysis. (d) Precision Analysis. (e) Recall Analysis.

V. CONCLUSION

In this work, the different machine learning algorithms is evaluated for the prediction of liver disease. Due to the subtle nature of its symptoms, liver disease is particularly difficult to diagnose. Liver disease prediction followed the step of preparing data in that data was collected from the public database preprocessing of data for -1 value replacement. Data division into the entire array of data split into training and research. Eventually, quantitative measurement metrics such as precision, accuracy and recall are measured over various machine learning models.

REFERENCES

1. Vasan Durai, Dinesh, Kalthireddy, "Liver disease prediction using machine learning", International Journal of Advance Research, Ideas and Innovations in Technology, 2017
2. Nazmun Nahar and Ferdous Ara, "Liver disease prediction by using different Decision tree techniques", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.8, No.2, March 2018
3. V.V. Ramalingam, A.Pandian, R. Ragavendran, "Machine Learning Techniques on Liver Disease A Survey", International Journal of Engineering & Technology, 7 (4.19) (2018) 485-495
4. Joel Jacob, Joseph Chakkalakal Mathew, Johns Mathew, Elizabeth Issac, "Diagnosis of Liver Disease Using Machine Learning Techniques" International Research Journal of Engineering and Technology (2018)
5. Shambel Kefelegn, Pooja Kamat, "Prediction and Analysis of Liver Disorder Diseases by using Data Mining Technique: Survey", International Journal of Pure and Applied Mathematics 2018
6. L. Alice Auxilia, "Accuracy Prediction using Machine Learning Techniques for Indian Patient Liver Disease", Proceedings of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018)
7. Ain Najwa Arbain, B. Yushalinie Pillay Balakrishnan, "A Comparison of Data Mining Algorithms for Liver Disease Prediction on Imbalanced Data", International Journal of Data Science and Advanced Analytics, 2018
8. G. Ignisha Rajathi 1,* and G. Wiselin Jiji 2, "Chronic Liver Disease Classification Using Hybrid Whale Optimization with Simulated Annealing and Ensemble Classifier", MDPI Journal 2019.
9. Hemanth Kumar and Sivasangari.A (2016), " An Efficient Distributed Data processing Method for Smart environment" Indian Journal of Science and Technology, ISSN (Print) : 0974-6846, Vol. 9, Issue 31,pp. 1-4.
10. P.Ajitha, A. Sivasangari, K.Indira,"Predictive Inter and Intra Parking System",International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, Vol.8, Issue-2S, December 2018.
11. R. M. Gomathi, G.H.S.Krishna, Brumancia.E, Y.Mistica Dhas,"A Survey on IoT Technologies, Evolution and Architecture", 2nd International Conference on Computer, Communication, and Signal Processing: Special Focus on Technology and Innovation for Smart Environment(ICCCSP 2018),IEEE. Feb. 2018.
12. Indira K, Christal Joy E, "Energy Efficient IDS for Cluster-Based VANETS",Asian Journal of Information Technology, vol 14(1) ,37-41,2015.
13. E. Brumancia, S. Justin Samuel, R. M. Gomathi and Y. Mistica Dhas,"An Effective Study On Data Fusion Models in Wireless Sensor Networks", ARPN Journal of Engineering and Applied Sciences, Vol. 13, No. 2, January 2018.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:36:32 UTC from IEEE Xplore. Restrictions apply.

International Conference on Communication and Signal Processing, July 28 - 30, 2020, India

Design of Mobile Robot Teleportation System using Virtual Reality

Sourav Ray and A. Sivasangari

***Abstract*—Home gadgets. This's really simple to home automation entry. In this paper, the design is presented by us and also setup associated with an inexpensive but versatile, doable and also protected virtual reality dependent house hands-free operation process. From our suggested program, we are going to try to manage lights, 230v supply, and motor. In order to resolve the issue of the delay time period that is unforeseen, immediate teleportation structure is suggested. This particular structure permits us to lessen the trajectory blunder on the equipment management that is certainly managed through the internet. This particular system performs additionally shows the usage of virtual reality within the context on equipment management. Although also within parallel together with the actual phone system to enhance quality management, virtual reality could be utilized inside a standard fashion to mimic the actions of a method using Robot Communication. In order to verify the work of ours, we conducted experimental tests in different locations. The usefulness of the suggested structure is shown by experimental outcomes.**

***Index Terms*—Robot, home automation, Arduino, Sensors, Mobile application, virtual reality**

I. Introduction

PRESENT trends and technologies are making developing as well as setup of lodged methods for the household automation operation has grown the skills as well as functions. There's a need for sensible home automation operation entry through virtual reality. The suggested method enables the users to get into through virtual reality in simple fact as well as balancing the principles powering are based on theories regarding much-held man wish to get away from the borders of the' real world' by adopting cyberspace.

As soon as generally there we are able to have interaction on these virtual surroundings within a far more naturalistic fashion that is going to generate brand new types of human-machine interaction (HMI). The target is moving past, common kinds of interaction like the computer keyboard and also computer mouse that a lot of people deal with daily. This's viewed as a bad method of dealing that causes customers to adjust to the needs instead of the opposite manner available. Though the alternative is done by a virtual atmosphere.

Sourav Ray and A. Sivasangari are with the Sathyabama Institute of Science and Technology, (e-mail: sivasangarikavya@gmail.com, saruve910@gmail.com).

Someone is allowed by it to completely immerse themselves inside a very visible environment that they explore byways of the senses of theirs. Home automation operation is one of several main uses that is appearing. A home automation operation device is applied by using Bluetooth functions. The frequency of it is 2.4GHz. It is able to relate products inside 10m to 100m in an acceleration of 3Mbps. With this particular ability of Bluetooth, a house automation operation device according to virtual reality is proposed by us.

The rest of the paper is organized as follows. Section II, describes the work related to the field. The proposed work is explained in section III. In section IV simulation results are presented and finally conclusion in section V.

II. Related Work

Some of the research papers are surveyed as per the area we took to develop and to design the system stronger. The development approach of ours is grounded on sent out symbolic processing via executable code exchange among devices, and also permits to lengthen the skills even with the deployment [1] [2] [3] of theirs.

Virtual reality may affect the real time action accuracy on comparing but if it is built with unique technologies connected then it will be the best out of all in the market, this paper tells about the system which connected with the Internet of things for faster communication and constant computational time [4]. This paper supplies introduction as well as distinct evaluation on the IoT structure paradigm with all the explanation of the fundamental requirements of its together with the setup problems as well as succeeding paths [5]. Survey on home automation sensors overseeing methods utilizing

Wireless sensor Networks is offered [6]. Present a ground for talking about available study issues in home automation using virtual reality, a perspective for exactly how IoT might alter the system within the distant long term is for starters offered. Next, 8 important studies subjects are enumerated as well as investigation issues inside these subject areas are talked about [7]. Provides the different strategies and also algorithms used within the style of committed virtual reality methods through the WSN [8].

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:38:32 UTC from IEEE Xplore. Restrictions apply.

Demonstrate a functionality evaluation around the usage of robot requirements within smartphone apps to ingest as well as generate environmentally friendly sensor information, analyzing to what degree the overall performance issues associated with XML may be relieved by utilizing option uncompressed as well as compressed platforms [9]. With this paper, 3 automation programs of wireless sensor system for example things theft management in home, home automation management, light Control are described in information.

At present contained advancement phase included in a bigger quality of the air overseeing technique task, the application program is going to enable subscribers to keep an eye on their day contact with gasoline contaminants by blending operator area information even some papers are saying these system will reduce the fire accident inside the home, also gave suggestion for home automation system and this paper tell about bi-directional communication approach between an IoT system. and a virtual avatar character 2 provides a background on IoT and HCI needs, dashboards, and mixed reality for IoT [10]. Out of all scientist suggestions, we have decided to propose the system using virtual reality technology to home automation with good computational time and communication speed.

III. Proposed Approach

Virtual Reality chip board built to encounter fresh types of human machine interaction (HMI). The benefit which make it possible for the inserted phone system custom to effortlessly, seamlessly and quickly pour together with the microcontroller by UART to the applications of theirs. It will interface with Laptop or PC near the USB offered. By operating the place program in application, you are able to set the distance by looking at the digital camera picture obtainable in the Laptop or PC in that the optimize interaction is able to accomplished.

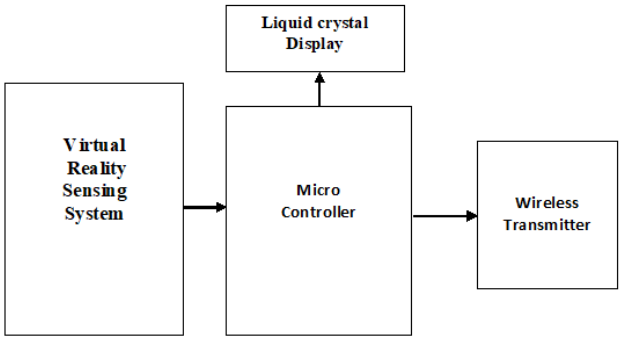


Fig. 1. Transmitter Diagram.

In Fig. 1 and Fig. 2, diagram is having electrical power product, LCD, Transmitter, receiver, ARM7 controller, and modem. Virtual reality device with Arduino analog pins will connect the entire system into one. Instead of using device 1 and device 2, we are using robots. The controller which in turn approach the device information worth as well as post to mobile app and server room. Mobile app is utilized to individuals for checking progress and connectivity problems., this particular system was created that can help an individual to identify, computer monitor, and then examination the automation inside a certain place.

A. Virtual Reality Transmitter

It will transmit the input to the controller, so that controller will bring the command through the wireless technology to the receiver. This system will help the user for operating the whole atmosphere inside the home devices in automation.

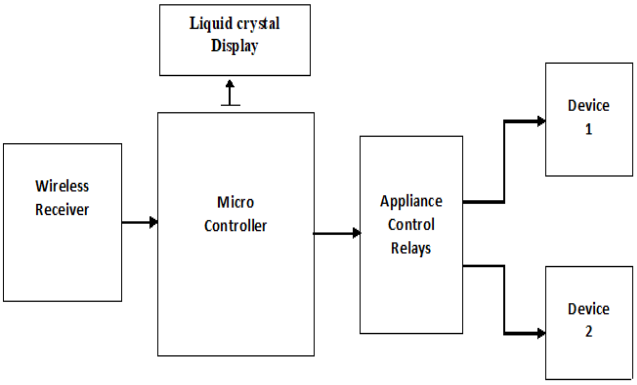


Fig. 2. Destination Diagram.

There are more types of virtual reality devices are there, we can use those as per the usage limit. We have tested and fits to this system.

B. Setting up Screen of appliance control

Screen was placed to check the whole atmosphere by sitting in the same place and detecting the actions of the devices. This is also useful to check the device status as well as availability. Appliance control will allow us to connect or disconnect the devices.

C. Sensing of end user and Wireless Transmission

User will send the command from virtual reality through wireless transmission to the receiver. Then the command to the receiver controller to the device consider as robot to operate the automation and bringing it to the next level. In Fig. 1, we have considered the devices as robot in our system.

D. Robot and Mobile Application

Lately, the benefits of smartphone have been changed by advancements in deep engineering. Robot will get the command from the user and perform as per the request to the home automation devices. Currently the electric market place is acquired by Android technology. The ROBOT along with regular devices to offer home information and device information. In case a person is placed away to a spot, then a warning is shown to the user through application. In case the problem amount is absurdly large therefore the end user is able to reroute the system as they want.

IV. Experimental Results

The experiments are performed using the ARDUINO IDE and Embedded C version. The computations are performed using virtual reality data that is readily available in chip board.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:38:32 UTC from IEEE Xplore. Restrictions apply.

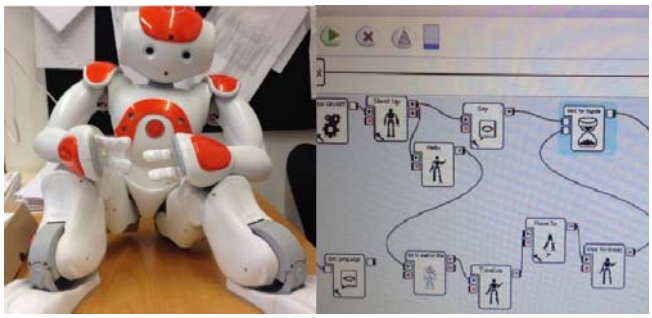


Fig. 3. Device Setup.



Fig. 4. Virtual reality device.

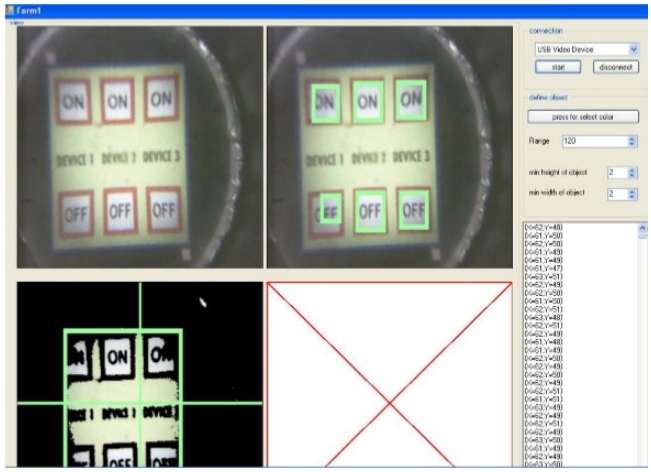


Fig. 5. Output Screenshot.

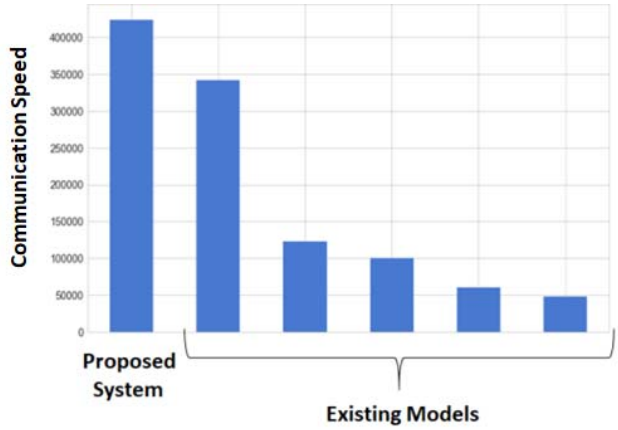


Fig. 6. Communication Speed.

In Fig. 3, Device setup, all devices are connected with the Display, Chip and controller are connected with the board for testing the response in many directions. Fig. 4 shows the virtual reality for accessing the device details and PC storage. In Fig. 5, we can have tested the home automation devices using virtual reality technology, here ON or OFF test was done to the particular area. Fig. 6 shows the communication level. The data are then trained with a proposed scheme which is widely used for all techniques. Some database is kept for training and the rest are kept for testing the proposed schemes. Hence the result satisfies the expected output, achieved the communication level on comparing with the existing model.

V. Conclusion

This Proposed scheme together with the incorporated movable virtual reality could be beneficial to individuals struggling with home automation problem. The system had using characteristics, indices of devices for a particular community utilizing real time computation, timing outside tasks for various suggestion. There more places are consider for testing this system and made in depth analysis for bring the accuracy and communication in stable. As we tested, communication level is high. One particular remedy could possibly be making use of Cyber Physical System (CPS), rather than IoT to lessen computation to improve the overall performance on the product. Here we have achieved the computational response and Communication speed on comparing with existing system.

References

[1] Callaghan, M. J., Bengloan, G., Ferrer, J., Cherel, L., El Mostadi, M. A., Eguíluz, A. G., & McShane, N. (2018, March). Voice driven virtual assistant tutor in virtual reality for electronic engineering remote laboratories. In International Conference on Remote Engineering and Virtual Instrumentation (pp. 570-580). Springer, Cham.

[2] de Tommaso, M., et al (2016). Testing a novel method for improving wayfinding by means of a P3b Virtual Reality Visual Paradigm in normal aging. Springerplus, 5(1), 1-12..

[3] Becker, G. (2018). U.S. Patent Application No. 16/059,291.

[4] Kepuska, V., & Bohouta, G. (2018, January). Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home). In 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC) (pp. 99-103).

[5] Vishwakarma, S. K., et al 2019 Smart energy efficient home automation system using iot. In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) (pp. 1-4). IEEE.

[6] Madakam, S., Holmukhe, R. M., & Jaiswal, D. K. (2019). The future digital work force: robotic process automation (RPA). JISTEM-Journal of Information Systems and Technology Management, 16.

[7] Choi, Y. K., Lazar, A., Demiris, G., & Thompson, H. J. (2019). Emerging smart home technologies to facilitate engaging with aging. Journal of gerontological nursing, 45(12), 41-48.

[8] Miori, V., Russo, D., & Ferrucci, L. (2019, October). Interoperability of home automation systems as a critical challenge for IoT. In 2019 4th International Conference on Computing, Communications and Security (ICCCS) (pp. 1-7). IEEE.

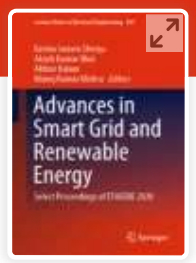
[9] Blanco-Novoa, Ó., et al 2019, Towards the Internet of Augmented Things: an Open-source Framework to Interconnect IoT Devices and Augmented Reality Systems. In Presented at the 6th International Electronic Conference on Sensors and Applications (Vol. 15, p. 30).

[10] Shao, Y., Lessio, N., & Morris, A. (2019). IoT Avatars: Mixed Reality Hybrid Objects for CoRe Ambient Intelligent Environments. Procedia Computer Science, 155, 433-440.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 04:38:32 UTC from IEEE Xplore. Restrictions apply.

Springer Link

Search Log in



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 509–517

Dengue Prediction Using Machine Learning Techniques

O. Shireesha, P. Hema Sri, T. Prem Jacob, G. Nagarajan & A. Pravin

Conference paper | First Online: 05 January 2021

304 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Dengue fever is a gigantic general medical problem that causes mortality and severity worldwide. As of now, dengue fever is the most vital irresistible disease. Postponing recognition of dengue disease could cause dangerous difficulties and increase the death rate. In this way, exploring aims to build an exact model that could more likely identify the first signs and side effects of dengue fever and a commonsense framework for disease self-defense. These two techniques have been applied to promptly notify patients, whether they are

suspected of having dengue fever or not specifically the data extraction method. Dengue detection rules are developed after meeting a professional, and these rules will be applied in an authority system using a flexible basis. Regardless of before applying the removed standards, the data mining gadget attempted the accuracy of the rules. Preliminary results show that data extraction has improved accuracy and produces a solid result for self-investigation of dengue marks.

Keywords

Dengue genetics **Severe dengue**

Machine learning

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_50
- Chapter length: 9 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Softcover Book EUR 279.99

Hardcover Book EUR 279.99

Learn about institutional subscriptions

References

1. P. Manivannan, P. Isakki, Dengue fever prediction using *K*-means clustering algorithm, in *IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing*, 978-1-5090-4778-9/17/©2017

2. I.I. Wiratmadja, S.Y. Salamah, R. Govindaraju, J. Eng. Technol. Sci. **50**(1), 110–126 (2018)

3. M.A. Chowdary, M. Kundan, D.A.V.A. Mary, Effective credit card forgery prevention using multilevel authentication. IOP Conf. Ser. Mater. Sci. Eng. 590(1), 012021 (2019) (IOP Publishing)

4. P. Mekha, K. Osathanunkul, et al., in *Gene Classification of Dengue Virus Type Based on Codon Usage*, 978-1-5090-4420-7/16/ ©2016 IEEE

5. T. Anandhi, K.V. Kishore, G.S. Maha, R.M. Gomathi, A sustainable vehicle parking using IoT, in *2019 3rd International Conference on Trends in Electronics and Informatics* (*ICOEI*). IEEE, Apr 2019, pp. 950–952

6. A.M. Najar, M.I. Irawan, et al., Extreme learning machine method for dengue hemorrhagic fever

outbreak risk level prediction, in *IEEE International Conference on Smart Computing and Electronic Enterprise* (*ICSCEE2018*) ©2018

7. M. Mufli Muzakki, F. Nhita, The spreading prediction of dengue hemorrhagic fever (DHF), in Bandung regency using *K*-means clustering and support vector machine algorithm, in *6th International Conference on Information and Communication Technology* (2018)

8. K.E. Gnanesh, T. Dheeraj Bhavan Narayana, M.D. Kamalesh, Retrieval of encrypted data using trapdoor method in cloud computing. J. Comput. Theoret. Nanosci. **16**(8), 3237–3241 (2019)

9. S.P. Shyry, Biometric-based three-tier microservice architecture for mitigating the fraudulent behaviour, in *Proceedings of the 2nd International Conference on Communication, Devices and Computing* (Springer, Singapore, 2020), pp. 399–404

10. R.S.B. Krishna, M. Aramudhan, Feature selection based on information theory for pattern classification, in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies* (*ICCICCT*). IEEE, July 2014, pp. 1233–1236

11. N. Srinivasan, C. Lakshmi, Stock price prediction using rule based genetic algorithm approach. Res. J. Pharm. Technol. **10**(1), 87–90 (2017)

12. P.S. Songkok, H.A. Wibawa, et al., Performance comparison of artificial neural network models for dengue fever disease detection, in *1st International Conference on Informatics and Computational Sciences* (*ICICoS*) (2017)

13. S.C. Mana, A feature based comparison study of big data scheduling algorithms, in *2018 International Conference on Computer, Communication, and Signal Processing* (*ICCCSP*) (Chennai, 2018), pp. 1–3

14. B. Omkar, D. Preet, et al., Int. Res. J. Eng. Technol. (IRJET) 04(10) (2017)

15. R. Yogitha, G. Mathivanan, Performance analysis of transfer functions in an artificial neural network, in *2018 International Conference on Communication and Signal Processing* (*ICCSP*). IEEE, Apr 2018, pp. 0393–0397

16. N.M. Zainee, K. Chellappan, A preliminary dengue fever prediction model based on vital

signs and blood profile, in *IEEE EMBS Conference on Biomedical Engineering and Sciences* (*IECBES*) (2016)

17. K. Sangeetha, P. Vishnuraja, D. Deepa, Stable clustered topology and secured routing using mobile agents in mobile ad hoc networks. Asian J. Inf. Technol. **15**(23), 4806–4811 (2016)

18. N. Mathur, V.S. Asirvadam, Visualization of dengue incidences using expectation-maximization (EM) algorithm. IEEE 978-1-5090-0845-2/16/© 2016

19. T. Marimuthu, V. Balamurugan, A novel bio-computational model for mining the dengue gene sequences. Int. J. Comput. Eng. Technol. (IJCET) **6**(10), 17–33 (2015). Article ID: IJCET_06_10_003

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

O. Shireesha, P. Hema Sri, T. Prem Jacob, G. Nagarajan & A. Pravin

Corresponding author

Correspondence to O. Shireesha .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Shireesha, O., Sri, P.H., Prem Jacob, T., Nagarajan, G., Pravin, A. (2021). Dengue Prediction Using Machine Learning Techniques. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in

Electrical Engineering, vol 691. Springer, Singapore.
https://doi.org/10.1007/978-981-15-7511-2_50

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7511-2_50

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

MENU SIGN IN/REGISTER

Scilight Summaries of the latest breakthroughs in the physical sciences AIP Publishing

AIP Publishing Conference Proceedings

HOME BROWSE MORE

Home > AIP Conference Proceedings > Volume 2207, Issue 1 > 10.1063/5.0000431

PREV NEXT

No Access

Published Online: 28 February 2020

Deep feature fusion for IRIS based on industrial biometric engineering

AIP Conference Proceedings **2207**, 040003 (2020); https://doi.org/10.1063/5.0000431

I. Surendra[1,a)], T. Sai Sashank[1,b)], M. D. Anto Praveena[1,c)], *and* R. Joseph Manoj[2,d)]

View Affiliations
View Contributors

PDF | E-READER

ABSTRACT

Students attendance is one of the primary and most vital things that is used in all the schools. Schools have numerous systems for taking attendance. Some of them use manual attendance, while some of them taking attendance by making use of the bar codes embedded in the student identity cards. Though various systems are been used still it needs ample development to make it more efficient with minimal time consumption. In this paper, we have proposed a model for taking the attendance of the students by making use of biometric attendance systems that are intended to make use of WFT (Wireless Fingerprint System). The system proposed in the paper makes use of the unique Aadhaar number that is given to all the citizens of India. The student biometric is matched with their respective Aadhaar number. By making use of the proposed system it is not only useful for attendance tracking but also useful for the government to provide scholarships or education fees waiver to the deserved students. The system makes use of local normalization, second stage enhancement and matching modules to make it more efficient when compared to other traditional attendance systems.

SELECT YOUR ACCESS

INDIVIDUAL ACCESS

PDF | E-READER

If you have an individual subscription, a subscription provided by one of AIP's Member Societies, have claimed access to a Conference Proceeding, or have made an individual purchase, sign in below.

Username:

Password

Remember me

LOG IN

Forgot password?

INSTITUTIONAL ACCESS

Access through your institution

PURCHASE

Standard PPV for $40.00

ADD TO CART



Resources

AUTHOR

PDF | E-READER

LIBRARIAN

ADVERTISER

General Information

ABOUT

CONTACT

HELP

PRIVACY POLICY

TERMS OF USE

FOLLOW AIP PUBLISHING:

Website © 2023 AIP Publishing LLC.
Article copyright remains as
specified within the article.

Scitation

PDF | E-READER

Springer Link

Search Log in



Fog Computing for Healthcare 4.0 Environments pp 337–354

Data Security and Privacy Functions in Fog Computing for Healthcare 4.0

A. Sivasangari, P. Ajitha, E. Brumancia, L. Sujihelen & G. Rajesh

Chapter | First Online: 03 August 2020

561 Accesses | **2** Citations

Part of the Signals and Communication Technology book series (SCT)

Abstract

Sensors play an essential role in different applications such as medicine, manufacturing, climate, smart transportation, and smart city. Wearable or implantable body sensors are necessary for the human body to collect patient information. Such tools produce a massive amount of data, and to collect useful information, it is more difficult to secure such data from intruders, process, and interpret it. In this chapter, we are improving such a big data health monitoring system by leveraging the fog computing principle at smart gateways, offering advanced network edge techniques and services. In particular, as a case

study, we chose electrocardiogram (ECG), because it plays an important role in the diagnosis of many heart diseases. The experimental results show that fog computing helps to reduce the encrypt and decrypt time compared to other traditional algorithms, and the information will be transmitted more safely using the algorithm with less computational overhead.

Keywords

Enhanced TEA algorithm

Optimized fully encryption algorithm

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-3-030-46197-3_14
- Chapter length: 18 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 117.69

Softcover Book EUR 149.99

Hardcover Book EUR 149.99

Learn about institutional subscriptions

References

1. Lu, R., Heung, K., Lashkari, A. H., Ghorbani, A. A. (2017). A Lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT. *IEEE Access Special Section On Security and Privacy In Applications and Services For Future Internet of Things, 5*, 3302–3312.

2. Liu, X., Yang, Y., Choo, K.-K. R., & Wang, H. (2018). Security and privacy challenges for internet-of-things and fog computing. *Hindawi Wireless Communications and Mobile Computing, 2018*, 1–3.

3. Zhang, X., Yuan, Y., Zhou, Z., Li, S., Qi, L., & Puthal, D. (2019). Intrusion detection and prevention in cloud, fog, and internet of things. *Hindawi Security and Communication Networks., 2019*, 1–4.

4. Aliyua F., Sheltamia, T., Shakshukib, E. M. (2018). *A detection and prevention technique for man in the middle attack in fog computing*. The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks EUSPN, pp. 24–31.

5. Lee, K., Kimy, D., Ha, D., Rajput, U., Oh, H. (2015). *On security and privacy issues of fog computing supported internet of things environment*. 6th

International Conference on the Network of the Future (NOF), IEEE.

6. Diro, A. A., Chilamkurti, N., & Kumar, N. (2017). *Lightweight cybersecurity schemes using elliptic curve cryptography* (Publish-Subscribe fog Computing) (Vol. 22, pp. 848–858). New York: Springer Science Business Media.

7. Zhu, J., Hu, J., Zhang, M., Chen, Y., & Bi, S. (2019). *A fog computing mode for implementing motion guide to visually impaired: Simulation modelling practice and theory*. New York: Science Direct.

8. Diro, A. A., Chilamkurti, N., & Nam, Y. (2018). Analysis of lightweight encryption scheme for fog-to-things communication. *IEEE Access Special Section on Real-Time Edge Analytics for Big Data in Internet of Things., 6*, 26820–26830.

9. Wang, Q., Chen, D., Zhang, N., Ding, Z., & Qin, Z. (2017). PCP: A privacy-preserving content-based publish–subscribe scheme with differential privacy in fog computing. *IEEE Access, 5*(2017), 17962–17974.

10. Zhang, Y., Zhao, J., Dong, Z., Deng, K., Ren, F., Zheng, X., & Shu, J. (2018). Privacy-preserving data aggregation against false data injection attacks in fog computing. *Sensors., 13*, 1–16.

11. Jia, X., He, D., Kumar, N., & Raymond, K.-K. (2019). Authenticated key agreement scheme for fog-driven IoT healthcare system. *Wireless Networks Springer, 25*, 4737–4750.

12. Bhaskar, K., Jayashree R., Sathiyavathi R., Mary Gladence L., and Maria Anu V. (2017). A novel approach for securing data de-duplication methodology in hybrid cloud storage. In *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, pp. 1–5. IEEE

13. Gope, P. (2019). LAAP: Lightweight anonymous authentication protocol for D2D-Aided fog computing paradigm. *Computers & Security Science Direct, 86*, 223–237.

14. Xiao, X., Hou, X., Liu, C., & Li, Y. (2019). Quantitative analysis for capabilities of vehicular fog computing. *Information Sciences Science Direct, 501*, 742–760.

15. Raafat Aburukba, Mazin Ali Karar, Taha Landolsi, Khaled El-Fakih. (2019). Scheduling internet of things requests to minimize latency in hybrid Fog–Cloud computing. Future Generation Computer Systems, Science Direct.

16. Islam, S. U., Akhunzada, A., Boudjadar, J., Khattak, H. A., Din, I. U., & Rodrigues, J. J. P. C. (2019). Energy and performance aware fog computing: A case of DVFS and green renewable energy. *Future Generation Computer Systems, 101*, 1112–1121.

17. Pereira, J., Ricardo, L., Luís, M., Senna, C., & Sargento, S. (2019). Assessing the reliability of fog computing for smart mobility applications in VANETs. *Future Generation Computer Systems., 94*, 317–332.

18. Zhao, D., Sun, G., Liao, D., Xu, S., & Chang, V. (2019). Mobile-aware service function chain migration in cloud–fog computing. *Future Generation Computer Systems, 96*, 591–604.

19. Tuli, S., Mahmud, R., Tuli, S., & Buyya, R. (2019). Fog Bus. *A Blockchain-based lightweight framework for edge and fog computing. Journal of Systems and Software., 154*(22–36).

20. Indira, K., & Christal, J. E. (2015). Energy Efficient IDS for Cluster-Based VANETS. *Asian Journal of Information Technology., 14*(1), 37–41.

21. Siva Sangari, A., Manickam, J. M. L., & Gomathi, R. M. (2016). RC6 based security. *Wireless Body*

Area Network, 74, 31–34.

22. Ajitha, S. A., & Indira, K. (2018). Predictive inter and intra parking system. *International Journal of Engineering and Advanced Technology., 8*, 354–357.

23. Ajitha, N. H. (2019). An adaptive approach for dynamic resource allocation in cloud service. *International Journal of Control Theory and Applications., 9*(10), 4871–4878.

24. Gomathi, R. M., & Manickam, J. M. L. (2019). Energy efficient static node selection in underwater acoustic wireless sensor network. *Wireless Personal Communications, 107*(2), 709–727.

25. Gomathi, R. M., Martin Leo Manickam, J., & Sivasangari, A. (2016). A comparative study on routing strategies for underwater acoustic wireless sensor network. *Contemporary Engineering Sciences., 9*(1–4), 71–80.

26. Brumancia, E., Justin Samuel, S., Gomathi, M., & Mistica Dhas, Y. (2018). An effective study on data fusion models in wireless sensor networks. *ARPN Journal of Engineering and Applied Sciences., 13*, 686–692.

27. Brumancia, E., & Sylvia, A. (2015). A profile based scheme for security in clustered wireless sensor networks. *International Conference on Communications and Signal Processing (ICCSP)* pp. 2–5.

28. Kumari, A., Tanwar, S., Tyagi, S., & Kumar, N. (2018). Fog computing for healthcare 4.0 environment: Opportunities and challenges. *Computers & Electrical Engineering., 72*, 1–13.

29. Tanwar, S., Tyagi, S., & Kumar, N. (Eds.) (2019). Multimedia Big Data computing for IoT applications: Concepts, paradigms and solutions. *Intelligent Systems Reference Library* (pp. 1–425). Springer Nature Singapore Pte Ltd., Singapore.

30. Mittal, M., Tanwar, S., Agarwal, B., & Goyal, L. M. (Eds.) (2019). *Energy conservation for IoT Devices: Concepts, paradigms and solutions, studies in systems, decision and control* (pp. 1–356). In Preparation, Springer Nature Singapore Pte Ltd., Singapore.

31. Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G-enabled IoT for industrial automation: A systematic review. *Solutions, and Challenges: Mechanical Systems and Signal Processing, 135*, 1–19.

32. Vora, J., Tanwar, S., Verma, J. P., Tyagi, S., Kumar, N., Obaidat, M. S., & Rodrigues, Joel J. P. C. (2018). *BHEEM: A Blockchain-based framework for securing electronic health records*. IEEE Global Communications Conference (IEEE GLOBECOM-2018), Abu Dhabi, UAE, pp. 1–6.

33. Vora, J., Devmurari, P., Tanwar, S., Tyagi, S., Kumar, N., & Obaidat, M. S. (2018). Blind signatures based secured e-healthcare system. *International Conference on Computer, Information and Telecommunication Systems (IEEE CITS-2018), Colmar, France, 11-13*, 177–181.

34. Patel, D., Narmawala, Z., Tanwar, S., & Singh, P. K. (2018). A systematic review on scheduling public transport using IoT as tool. In B. Panigrahi, M. Trivedi, K. Mishra, S. Tiwari, & P. Singh (Eds.), *Smart innovations in communication and computational sciences* (Advances in Intelligent Systems and Computing) (Vol. 670, pp. 39–48). Singapore: Springer.

35. Sivasangari, A., Bhowal, S., & Subhashini, R. (2019). Secure encryption in wireless body sensor networks. *Emerging Technologies in Data Mining and Information Security, 3*, 679–686.

36. Kumari, A., Tanwar, S., Tyagi, S., & Kumar, N. (2019). Verification and validation techniques for streaming Big Data analytics in internet of things environment. *IET Networks, 8*, 155–163.

37. Kumari, A., Tanwar, S., Tyagi, S., Kumar, N., Maasberg, M., & Choo, K. K. R. (2018). Multimedia Big Data computing and internet of things applications: A taxonomy and process model. *Journal of Network and Computer Applications, 124*, 169–195.

38. Sujihelen, L., & Jayakumar, C. (2018). Inclusive elliptical curve cryptography (IECC) for wireless sensor network efficient operations. *Wireless Personal Communication, 99*, 893–914.

Author information

Authors and Affiliations

School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

A. Sivasangari, P. Ajitha, E. Brumancia & L. Sujihelen

Department of IT, MIT campus, Anna University, Chennai, India

G. Rajesh

Editor information

Editors and Affiliations

Department of Computer Science and Engineering, Institute of Technology, Nirma University, Ahmedabad, Gujarat, India

Dr. Sudeep Tanwar

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Switzerland AG

About this chapter

Cite this chapter

Sivasangari, A., Ajitha, P., Brumancia, E., Sujihelen, L., Rajesh, G. (2021). Data Security and Privacy Functions in Fog Computing for Healthcare 4.0. In: Tanwar, S. (eds) Fog Computing for Healthcare 4.0 Environments. Signals and Communication Technology. Springer, Cham. https://doi.org/10.1007/978-3-030-46197-3_14

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-46197-3_14

Published	Publisher Name	Print ISBN
03 August 2020	Springer, Cham	978-3-030-46196-6

Online ISBN	eBook Packages

978-3-030-46197-3

Computer Science

Computer Science

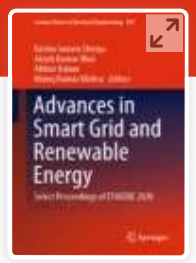
(R0)

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 491–498

Customer Stress Prediction in Telecom Industries Using Machine Learning

K. N. R. Srinivas ✉, K. S. S. Manikanta, T. Prem Jacob, G. Nagarajan & A. Pravin

Conference paper | First Online: 05 January 2021

308 Accesses | **1** Citations

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

This research work contains various types of developments and solutions to find customer stress prediction in the telecom industries to improve the facilities and supervised algorithms. In the competitive world, especially in the enterprise market, maintaining valuable customers is becoming a difficult task. In one situation, losing a customer is like a decrease in profits for telecom industry growth. In another situation, the price of acquiring new customers is much higher than the cost of retaining existing customers. For this critical

situation, the telecom industries should focus on keeping existing customers. This paper will analyze the customer data, which was collected as an open dataset, and predict customer stress by applying supervised machine learning algorithms mainly using linear discriminant analysis, support vector machine, and random forest.

Keywords

Customer stress **Support vector machine**

Random forest **Linear discriminant analysis**

One hot encoding

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_48
- Chapter length: 8 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33
Price includes VAT (India)

- ISBN: 978-981-15-7511-2
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 279.99
Price excludes VAT (India)

- ISBN: 978-981-15-7714-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 279.99
Price excludes VAT (India)

- ISBN: 978-981-15-7510-5
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. J.T. Tou, R.C. Gonzalez, *Pattern Recognition Principles* (Addison-Wesley, 1974)

2. D. Van den Poel, B. Larivière, Customer attrition analysis for financial services using proportional hazard models. Eur. J. Oper. Res. **157**, 196–217 (2004)

3. C.P. Wei, I.T. Chiu, Turning telecommunications call details to churn prediction: a data mining approach. Exp. Syst. Appl. **23**, 103–112 (2002)

4. G.V.K. Sai, P.S. Kumar, A.V.A. Mary, Incremental frequent mining human activity patterns for health care applications. IOP Conf. Ser. Mater. Sci. Eng. **590**(1), 012050 (IOP Publishing)

5. B. Bharathi, P. Vijayalakshmi, T. Nagarajan, Speaker identification using utterances correspond to speaker-specific-text, in *IEEE Technology Students' Symposium*. IEEE, Jan 2011, pp. 171–174

6. B.K. Wong, T.A. Bodnovich, Y. Selvi, Neural network applications in business: a review and analysis of the literature (1988–1995). Decis. Supp. Syst. **19**, 301–320 (1997)

7. S.M. Vaniya, B. Bharathi, Exploring object segmentation methods in visual surveillance for human activity recognition, in *2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication* (*ICGTSPICC*). IEEE, Dec 2016, pp. 520–525

8. X. Zhang, J. Edwards, J. Harding, Personalized online sales using web usage data mining. Comput. Ind. **58**, 772–782 (2007)

9. M.J.A. Berry, G.S. Linoff, *Data Mining Techniques: For Promotion Sales and Customer Support* (Wiley, 2003)

10. J. Sujolincy, M.D. Kamalesh, An authenticated smart card service using cross-cloud and big data, in *2016 International Conference on Control, Instrumentation, Communication and Computational Technologies* (*ICCICCT*). IEEE, Dec 2016, pp. 634–638

11. G. Zhan, B.E. Patuwo, M.Y. Hu, Forecasting with artificial neural network: the state of the art. Int. J. Forecast. **14**, 35–62 (1998)

12. R.S.B. Krishna, M. Aramudhan, Unsupervised spectral sparse regression feature selection using social media datasets, in *Proceedings of the International Conference on Informatics and Analytics*, Aug 2016, pp. 1–5

13. N. Srinivasan, C. Lakshmi, A novel prediction based tree structured data using machine learning techniques. Res. J. Pharm. Biol. Chem. Sci. **7**(5), 527–531 (2016)

14. M.V. Ishwarya, D. Deepa, S. Hemalatha, A. Venkata Sai Nynesh, A. Prudhvi Tej, Gridlock surveillance and management system. J.

Comput. Theoret. Nanosci. **16**(8), 3281–3284 (2019)

15. W. Buckinx, D. Van den Poel, Customer base studies partial defection of loyal behavioural clients in a non-contractual FMCG trade setting. Eur. Mag. Oper. Res. **164**(1), 252–268 (2005)

16. S.P. Shyry, Novel enhanced encryption algorithm for shared key generation, in *Proceedings of the 2014 International Conference on Interdisciplinary Advances in Applied Computing*, Oct 2014, pp. 1–7

17. J. Burez, D. Van den Poel, CRM at a pay-TV company: using systematic models to decrease customer grinding down by targeted marketing for subscription services. Exp. Syst. Appl. **32**, 277–288 (2007)

18. R. Aishwarya, R. Yogitha, V. Kiruthiga, Smart road surface monitoring with privacy preserved scheme for vehicle crowd sensing. J. Comput. Theor. Nanosci. **16**(8), 3204–3210 (2019)

19. D.A. Chiang, Y.F. Wang, S.L. Lee, C.J. Lin, A goal-oriented sequential pattern for network

banking churn analysis. Exp. Syst. Appl. **25**, 293–302 (2003)

20. K. Coussement, D. Van den Poel, Churn prediction in subscription services: an application of support vector machines while comparing two parameter-selection techniques. Exp. Syst. Appl. **34**, 313–327. H. Cyben, Approximation by superpositions of sigmoidal (1998)

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

K. N. R. Srinivas, K. S. S. Manikanta, T. Prem Jacob, G. Nagarajan & A. Pravin

Corresponding author

Correspondence to K. N. R. Srinivas .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia
Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India
Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Srinivas, K.N.R., Manikanta, K.S.S., Prem Jacob, T., Nagarajan, G., Pravin, A. (2021). Customer Stress Prediction in Telecom Industries Using Machine Learning. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_48

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7511-2_48

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

ScienceDirect

Corrosion Protection at the Nanoscale

Micro and Nano Technologies

2020, Pages 295-314

Chapter 16 - Corrosion behavior of ceramic nanocomposite coatings at nanoscale

A.M. Kamalan Kirubaharan, P. Kuppusami

Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Available online 20 March 2020, Version of Record 20 March 2020.

Show less

Outline | Share | Cite

https://doi.org/10.1016/B978-0-12-819359-4.00016-7

Get rights and content

Abstract

This chapter explores the advancement in the development of ceramic nanocomposite coatings with a major emphasis on corrosion protection of metals and alloys. Examples on corrosion behavior of metal–ceramic nanocomposite coatings of Ni–SiO_2, Ni–Al_2O_3, Ni–TiO_2, and Ni–YSZ on corrosion are briefly presented from the corrosion point of view. The various contributing factors for corrosion mechanism such as choice of elements, grain size, texture, porosity, grain boundary, and concentration of ceramic oxide in nanocomposites are described. In addition, the development of compositionally graded Ni–YSZ coating and its corrosion behavior are illustrated to bring out the importance of material selection to reduce thermal expansion and lattice mismatch for the high-temperature corrosion applications. The new strategies and approaches toward corrosion protection at nanoscale are proposed based on the experimental results reported in the literature.

Previous

Next

Keywords

Compositionally graded coatings; Corrosion; Diffusion barrier coatings; Metal–ceramic nanocomposites; Physical vapor deposition

Recommended articles

Cited by (4)

Metal-ceramic diffusion barrier nanocomposite coatings on nickel based superalloys for corrosion and high temperature oxidation resistance

2022, Ceramics International

Show abstract

Inter-diffusion effects and tribological behaviour of electron beam evaporated Ni-YSZ nanocomposite coatings subjected to diffusion annealing with borosilicate glass for nuclear applications

2022, Ceramics International

Show abstract

Nano-engineered Material and Remediation Strategy

2023, Environmental Footprints and Eco-Design of Products and Processes

State-of-the-art review of fabrication, application, and mechanical properties of functionally graded porous nanocomposite materials

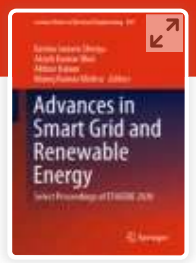
2022, Nanotechnology Reviews

Copyright © 2020 Elsevier Inc. All rights reserved.

Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

ELSEVIER

RELX™



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 555–561

Content-Based Image Retrieval in Cloud Image Repositories

Vishnu Vardhan Kathika, K. S. G. Jagadeesh Babu, A. Pravin, T. Prem Jacob & G. Nagarajan

Conference paper | First Online: 05 January 2021

301 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Cloud computing is a brand new kind of Internet-based computing which offers discussed laptop processing energy as well as information to other products and computers on need. It is the delivery of hosted products with the Web. Cloud computing services might be public, private, or even hybrid. The growing market of cloud has provide a service paradigm of storage/computation outsourcing will help to decrease user's burden of IT infrastructure maintenance and lower the price for both specific users and the enterprises. The three major

advantages of cloud computing are actually self service provisioning, elasticity, and pay per usage. The main purpose of this paper is to propose a secure frame work for outsourced privacy preserving storage and retrieval in large shared image repositories. In this paper, we had achieved the communication level when comparing with existing system and also achieved the accuracy level to the best.

Keywords

Cloud image **Accuracy** **Self service** **Encryption** **BOEW**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_55
- Chapter length: 7 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Softcover Book EUR 279.99

Hardcover Book EUR 279.99

Learn about institutional subscriptions

References

1. G. Nithya, S.L. Shabu, A novel framework in reusing the ontological health record. Res. J. Pharm. Biol. Chem. Sci. **7**(3), 215–220 (2016)

2. R.S.B. Krishna, M. Aramudhan, Feature selection based on information theory for pattern classification, in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies* (*ICCICCT*). IEEE, July 2014, pp. 1233–1236

3. P.S. Maran, P.M. Velumurugan, B.P.D. Batvari, Wind characteristics and Weibull parameter analysis to predict wind power potential along the South-East Coastline of Tamil Nadu, in *International Conference on Intelligent Information Technologies* (Springer, Singapore, 2018), pp. 190–199

4. A. Karthik, M.D. Kamalesh, Rat trap: inviting, detection & identification of attacker using honey words in purchase portal, in *2017 Third International Conference on Science Technology Engineering & Management* (*ICONSTEM*). IEEE, Mar 2017, pp. 130–132

5. S.C. Mana, B.K. Samhitha, J. Jose, M.V. Swaroop, P.C.K. Reddy, Traffic violation detection using principal component analysis and viola jones

algorithms. Int. J. Rec. Technol. Eng. (IJRTE) **8**(3) (2019). ISSN: 2277-3878

6. M.V. Ishwarya, D. Deepa, S. Hemalatha, A.V.S. Nynesh, A. Prudhvi Tej, Gridlock surveillance and management system. J. Comput. Theor. Nanosci. **16**(8), 3281–3284 (2019)

7. A. Ponraj, Optimistic virtual machine placement in cloud data centers using queuing approach. Future Gener. Comput. Syst. **93**, 338–344 (2019)

8. R. Yogitha, G. Mathivanan, Performance analysis of transfer functions in an artificial neural network, in *2018 International Conference on Communication and Signal Processing* (*ICCSP*). IEEE, Apr 2018, pp. 0393–0397

9. M. Selvi, P.M. Joe Prathap, Analysis & classification of secure data aggregation in wireless sensor networks. Int. J. Eng. Adv. Technol. **8**(4), 1404–1407 (2019)

10. N. Srinivasan, C. Lakshmi, Forecasting stock price using soft computing techniques, in *2017 Third International Conference on Science Technology Engineering & Management* (*ICONSTEM*). IEEE, Mar 2017, pp. 158–161

11. V.V. Kaveri, V. Maheswari, A model based resource recommender system on social tagging data. Indian J. Sci. Technol. **9**, 25 (2016)

12. R. Neha, J.Y. Bevish, A survey on security challenges and malicious vehicle detection in vehicular ad hoc networks. Contemp. Eng. Sci. **8**(5), 235–240 (2015)

13. S. Ancy, R. Kumar, R. Asokan, R. Subhashini, Prediction of onset of south west monsoon using multiple regression, in *Proceedings of IEEE International Conference on Computer Communication and Systems* (*ICCCS14*). IEEE, Feb 2014, pp. 170–175

14. M. Meeker, Internet trends 2015, in *Proceedings of the Code Conference* (2015), pp. 1–196

15. Global Web Index, Instagram tops the list of social network growth, 2013 [Online]. Available: https://tinyurl.com/hnwwlzm

16. N. Srinivasan, C. Lakshmi, Stock prediction and analysis using intermittent training data with artificial neural networks, in *2017 International Conference on Innovations in Information, Embedded and Communication Systems* (*ICIIECS*). IEEE Mar 2017, pp. 1–4

17. C.D. Manning, P. Raghavan, H. Schutze, *An Introduction to Information Retrieval*, vol. 1 (Cambridge University Press, Cambridge, UK, 2009)

18. S.P. Shyry, Efficient identification of bots by *K*-means clustering, in *Proceedings of the International Conference on Soft Computing Systems* (Springer, New Delhi, 2016), pp. 307–318

19. R. Chow, et al., Controlling data in the cloud: outsourcing computation without outsourcing control, in *Proceedings of the ACM Workshop Cloud Computing Security* (2009), pp. 85–90

20. A. Chen, GCreep: google engineer stalked teens, spied on chats (2010) [Online]. Available: https://gawker.com/5637234

21. D. Rushe, Google: don't expect privacy when sending to Gmail, 2013 [Online]. Available: https://tinyurl.com/kjga34x

22. G. Greenwald, E. MacAskill, NSA prism program taps into user data of Apple, Google and others (2013) [Online]. Available: https://tinyurl.com/oea3g8t

23. J. Halderman, S. Schoen, Lest we remember: cold-boot attacks on encryption keys. Commun. ACM **52**(5), 91–98 (2009)

24. National vulnerability database, in *CVE Statistics* (2014) [Online]. Available: https://web.nvd.nist.gov/view/vuln/statistics. Statistics (2014) [Online]. Available: https://web.nvd.nist.gov/view/vuln/statistics

25. D. Lewis, *iCloud Data Breach: Hacking and Celebrity Photos* (2014) [Online]. Available: https://tinyurl.com/nohznmr

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Vishnu Vardhan Kathika, K. S. G. Jagadeesh Babu, A. Pravin, T. Prem Jacob & G. Nagarajan

Corresponding author

Correspondence to Vishnu Vardhan Kathika .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Kathika, V.V., Jagadeesh Babu, K.S.G., Pravin, A., Prem Jacob, T., Nagarajan, G. (2021). Content-Based Image Retrieval in Cloud Image Repositories. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_55

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7511-2_55

Published	Publisher Name	Print ISBN
05 January 2021		

	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Comparing Machine-Learning Algorithms for Anticipating the Severity and Non-Severity of a Surveyed Bug

Muppala Sunny Chowdhary[1]
Student
chowdarysunny89@gmail.com
Akula Abinay[3]
Student
akulaabinay@gmail.com

R. Aishwarya[2]
Assistant Professor
aishwaryarajkumar7@gmail.com
Pasthupuram Harikrishna[4]
Student
masterkrishna1037@gmail.com

[1234]Sathyabama Institute of Science and Technology

Abstract: Programming Bug announcing is a vital piece of the product advancement process. When a Bug is accounted for on the Bug detection, their qualities are broken down and along these lines relegated to different fixers for their goals. Bug severity, a characteristic of a product bug survey is the level of effect that imperfection has on the advancement or activity of a segment or framework. In this paper, an endeavour has been made to show the pertinence of AI calculations to be specific SVM, KNN, Naïve Bayes, Naïve Bayes Multinomial, J48 and RIPPER in order to find out which classification algorithm is accurate to find the bug severity. The pertinence of the calculation in deciding the different degrees of bug severity for bug storehouses have been approved utilizing different execution measures by applying 5-overlap cross-approval.

Keywords Supervised Classification, Machine Learning, Bug Severity, Feature Selection, classification algorithms.

I. INTRODUCTION

Bug revealing and repairing is a significant period of programming improvement, refinement and upkeep for both open and shut source ventures. Different bug revealing frameworks have been produced for accommodation or detailing of a bug and keeping tabs on their development of fix. Bug revealing and following frameworks give a stage to record the issues/disappointments looked by the customer or client of the product. The direness of the bug needs a parameter that should be given during the bug announcing process. The crucial situation of the bug is severity of bug. There are 5 levels of bug survey severity. The synopsis and portrayal of the announced bug contain running abridged and nitty-gritty printed data about the bug. This printed data of the bug as of now submitted bug surveys having characterized severity levels is very helpful in deciding the severity level, need and task to the possibility fixer for the recently submitted bug survey utilizing directed arrangement.

II. PROBLEM STATMENTS

1. What classification algorithm would it be advisable for us to utilize when Anticipating bug survey severity?
2. What number of bug surveys are important when preparing a classifier based solid expectations?
3. What would we be able to find from the subsequent classification algorithms?

III. MATERIAL AND METHODOLOGY

The survey of bug data is made open detector models in software engineering (PROMISE). The values of bug have been surveyed shown in the table below. The below files are downloaded form CSV archieves. This methodology starts with demonstration, word termination removal, word hacting using the Doorman hacting figuring, and customer described once-over of stop words.

Projects	Severity wise Number of reported bugs						Number of terms
	1	2	3	4	5	total	
PitsA	0	111	155	34	3	303	697
PitsB	0	23	523	382	59	987	1001
PitsC	0	0	132	177	7	316	309
PitsD	0	1	167	13	1	182	347
PitsE	0	24	517	243	41	825	1059
PitsF	0	9	477	209	48	743	999

Table I. Severity in both no. of terms and bugs

Terms are additionally sifted through having a length under 3 and in excess of 50 characters. The bug surveys are spoken to utilizing a framework where each line speaks to the detailed bugs while every segment speaks to the stem of the terms/include and the last segment contains the seriousness level of the revealed bug.The proportion value denotes DG*JBG where the recurrence terms denotes DG and proportion logarithm denotes JBG.. In the present investigation, the following classification algorithms

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

are which are SVM, KNN, Naïve Bayes, Naïve Bayes Multinomial, J48 and RIPPER.

IV. CLASSIFER BUILIDING

Right now, we provide a glance at how bugs are accounted for and overseen inside a product venture. At that point, we give a presentation of information mining all in all and how we can utilize content mining with regards to bug surveys.

A. *The bug surveying and triaging*

Bug surveys exchange data between the clients of a product venture encountering bugs and the designers adjusting these flaws. Such a survey incorporates a one-line synopsis of the watched glitch and a more drawn out increasingly significant depiction, which may, for example, incorporate a stack follow. Normally the correspondent likewise includes data about the specific item and segment of the defective programming framework.

B. *Data mining and classification*

Most of the algorithm for classifications are concentrated in the field of data mining, every one of which acts contrastingly in a similar circumstance as per their particular qualities. Right now in this investigation, algorithm for classification such as SVM, KNN, Naïve Bayes, Naïve Bayes Multinomial, J48 and RIPPER on our concern to discover we get to a conclusion that which Algorithm is most appropriate for characterizing bug surveys in either a severe or a non-severe category.

C. *Preaction of a surveyed bug*

Each individual term happening in the records is viewed as an extra measurement, text information can be of large proportionality. Presenting preaction steps somewhat conquers this issue by decreasing the quantity of thought about terms.

Original description	Evolution crashes trying to open calendar
After stop-words removal	Evolution crashes open calendar
After stemming	evolut crash open calendar

Table II: Effects of the pre-processing steps

The typical preprocessing steps are Demonstration, Word termination removal, Hacting.

V. TESTING OF CLASSIFER

The precision of different AI systems for PITS ventures appears in Table III. From table III, the estimations of given below,

.

projects	No. of Terms	Accuracy of Machine Learning Techniques								
		NB	1-NN	2-NN	3-NN	4-NN	NBM	SVM	J4.8	RIPPER
PitsA	25	56.11	73.93	74.92	75.58	76.24	77.23	66.34	62.71	67.33
	50	60.73	74.59	77.23	75.58	79.21	76.57	65.68	66.34	66.01
	75	65.02	75.91	77.56	75.58	77.23	76.77	65.35	64.69	67.66
	100	68.98	73.93	77.23	75.25	75.91	77.23	66.34	68.32	67.00
	125	73.60	76.90	77.23	78.88	77.89	76.90	66.01	64.03	67.66
	150	76.24	76.24	77.23	77.56	77.89	77.23	67.33	63.37	64.69
	175	77.89	75.91	78.22	77.89	78.57	76.50	68.98	67.99	66.67
	200	78.88	76.57	77.89	79.21	78.55	77.56	69.31	68.65	64.69
PitsB	25	28.98	60.49	58.76	60.18	60.49	58.76	57.55	55.93	58.66
	50	38.60	64.74	62.92	64.64	64.54	62.11	60.28	57.24	59.78
	75	49.44	66.77	64.54	63.32	64.54	63.22	60.49	57.14	59.27
	100	51.17	65.25	63.22	66.46	66.46	64.03	60.79	58.36	59.78
	125	55.22	66.26	62.92	65.35	63.83	64.13	61.90	58.76	60.18
	150	58.05	66.36	63.73	64.84	65.35	64.44	62.72	58.87	60.08
	175	59.68	66.16	63.53	64.70	65.45	64.94	62.72	59.37	59.68
	200	60.59	66.06	64.44	65.86	65.45	65.05	62.82	60.28	59.88
PitsC	25	77.22	91.77	90.19	92.09	91.46	85.76	81.96	81.33	81.65
	50	86.71	90.19	88.29	88.29	87.97	85.76	81.33	83.86	77.85
	75	87.34	91.14	88.29	87.34	87.34	83.54	80.06	83.23	77.85
	100	87.03	91.14	87.66	87.66	86.71	84.18	82.28	83.23	79.11
	125	88.29	91.77	88.29	88.29	86.39	84.81	82.59	83.54	78.80
	150	89.56	91.14	87.66	87.66	85.76	86.08	84.81	83.86	73.73
	175	87.97	90.51	87.03	86.71	85.76	88.29	83.86	81.65	79.75
	200	86.71	89.24	85.44	85.44	84.49	89.24	86.08	80.38	76.58
PitsD	25	93.96	95.05	95.60	95.05	95.05	91.76	96.70	96.15	94.51
	50	92.31	93.41	95.60	95.05	95.60	91.76	96.15	95.60	95.05
	75	94.51	93.96	95.60	95.60	95.60	91.76	96.15	95.60	93.96
	100	92.86	93.41	95.60	95.60	95.60	91.76	96.15	95.60	93.96
	125	89.56	93.96	95.05	95.05	95.60	92.31	96.15	95.60	95.05
	150	88.46	94.51	95.05	95.05	95.60	92.86	96.15	95.60	93.41
	175	86.26	94.51	95.05	95.05	95.60	93.41	96.15	95.60	93.96
	200	87.36	94.51	95.05	95.05	95.60	93.41	96.15	95.60	92.86
PitsE	25	28.85	65.33	65.58	66.79	67.03	65.33	69.45	70.30	68.97
	50	33.70	64.97	65.82	65.58	65.58	66.79	69.58	70.42	68.97
	75	38.30	66.30	67.39	65.70	68.00	67.39	69.09	70.42	69.21
	100	43.88	68.97	67.52	68.48	70.30	67.76	69.45	69.94	69.21
	125	43.64	67.88	67.88	70.18	69.21	68.61	69.82	69.94	68.97
	150	48.24	69.09	70.30	70.67	70.67	68.73	69.33	70.06	69.21
	175	49.70	69.70	70.06	71.52	71.88	69.45	69.94	70.06	70.06
	200	51.64	70.42	71.64	73.09	73.09	69.09	70.79	69.82	69.94
PitsF	25	42.34	68.51	71.20	68.78	70.79	65.33	69.89	67.34	66.80
	50	47.45	69.31	74.16	72.68	73.35	66.79	71.24	67.34	67.07
	75	49.19	67.83	72.54	71.20	73.22	67.39	72.58	67.34	66.80
	100	51.75	68.64	72.81	70.93	72.41	67.76	72.45	67.34	66.94
	125	52.42	69.31	73.22	73.62	74.02	68.61	72.31	68.55	68.41
	150	55.24	69.58	72.41	72.54	71.87	68.73	72.45	69.49	69.89
	175	56.72	69.85	73.89	73.62	74.02	69.45	73.92	70.03	70.16
	200	59.01	71.06	75.10	74.29	73.89	69.09	72.98	69.89	70.83

Table III. Existing machine learning techniques accuracy in different no. of terms.

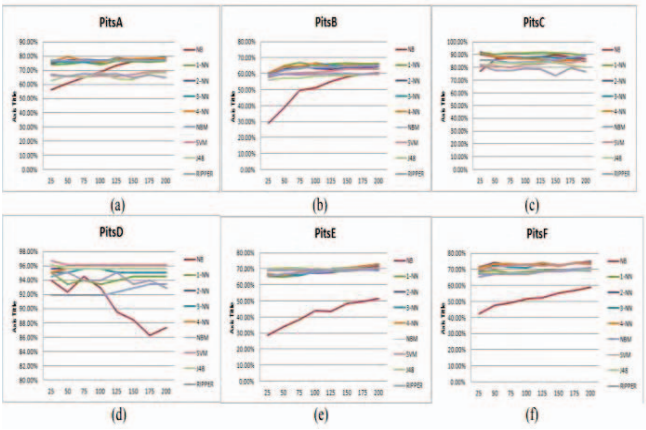


Fig 1. Concern graphs for the above tabular values are shown below for the pits A, pits B, pits C, pits D, pits E, pits F.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

Projects →		PitsA		PitsB		PitsC		PitsD		PitsE		PitsF	
Algorithm	*Severity*	*Maximum*	*Minimum*	*Maximum*	*Minimum*	*Maximum*	*Minimum*	*Maximum*	*Minimum*	*Maximum*	*Minimum*	*Maximum*	*Minimum*
NB	pred. 3	80.98%	51.95%	74.19%	44.73%	89.58%	85.26%	96.99%	92.26%	65.03%	44.10%	66.63%	53.54%
	pred. 2	87.62%	83.87%	37.04%	21.06%	-	-	-	-	12.38%	7.49%	21.05%	8.89%
	pred. 5	-	-	33.33%	14.74%	42.86%	16.66%	-	-	26.15%	19.81%	40.46%	27.66%
	pred. 4	43.75%	28.23%	52.28%	21.17%	91.41%	82.25%	73.34%	41.03%	48.36%	30.42%	56.12%	41.96%
1-NN	pred. 3	78.57%	74.83%	75.87%	72.61%	91.41%	82.25%	97.33%	96.41%	79.77%	76.26%	79.56%	77.84%
	pred. 2	87.89%	83.40%	53.85%	47.27%	-	-	-	-	20.83%	9.52%	66.67%	53.33%
	pred. 5	-	-	34.71%	14.89%	71.43%	62.50%	-	-	57.50%	32.84%	44.90%	14.81%
	pred. 4	46.38%	28.07%	63.17%	42.03%	93.37%	90.75%	74.07%	66.67%	57.68%	36.16%	59.13%	43.21%
2-NN	pred. 3	82.41%	77.76%	76.64%	71.94%	89.93%	85.71%	97.63%	97.33%	82.34%	77.93%	83.75%	81.67%
	pred. 2	86.79%	79.93%	45.46%	32.00%	-	-	-	-	19.61%	14.29%	62.50%	50.00%
	pred. 5	-	-	31.54%	12.50%	46.16%	40.00%	-	-	50.00%	28.95%	48.42%	16.44%
	pred. 4	28.00%	5.53%	51.76%	34.84%	92.08%	87.04%	75.00%	72.00%	49.73%	30.32%	56.97%	44.07%
3-NN	pred. 3	81.48%	78.05%	74.52%	71.43%	91.05%	85.30%	97.63%	97.33%	82.16%	76.98%	81.87%	79.56%
	pred. 2	89.69%	85.72%	32.26%	23.26%	-	-	-	-	24.39%	6.46%	62.50%	57.14%
	pred. 5	-	-	33.33%	6.81%	76.92%	20.00%	-	-	47.76%	28.98%	50.00%	20.69%
	pred. 4	33.90%	12.24%	60.63%	45.75%	93.37%	87.79%	75.00%	74.07%	58.83%	40.11%	60.76%	43.51%
4-NN	pred. 3	83.08%	78.66%	76.50%	72.00%	90.49%	85.22%	97.63%	97.33%	82.41%	78.09%	82.95%	81.12%
	pred. 2	88.99%	84.51%	35.50%	24.69%	-	-	-	-	27.78%	6.90%	62.50%	53.33%
	pred. 5	-	-	32.00%	11.36%	76.92%	20.00%	-	-	47.24%	30.00%	49.49%	19.67%
	pred. 4	28.00%	4.65%	57.63%	44.29%	92.70%	85.97%	76.92%	76.92%	57.69%	37.43%	58.38%	46.88%
NBM	pred. 3	81.50%	80.55%	75.63%	72.49%	89.06%	80.34%	96.53%	95.70%	80.84%	78.72%	80.84%	78.72%
	Pred. 2	84.88%	83.00%	-	-	-	-	-	-	-	-	-	-
	pred. 5	-	-	-	-	-	-	-	-	4.76%	4.76%	4.76%	4.76%
	pred. 4	-	-	54.12%	30.77%	91.06%	86.96%	88.66%	14.28%	41.50%	16.80%	41.50%	16.80%
SVM	pred. 3	76.81%	74.57%	74.23%	71.69%	82.46%	73.93%	96.99%	92.26%	81.56%	80.29%	83.30%	80.89%
	pred. 2	65.46%	55.85%	68.42%	53.33%	-	-	-	-	15.38%	13.79%	61.53%	57.14%
	pred. 5	-	-	12.70%	6.35%	85.71%	40.00%	-	-	54.25%	32.65%	25.45%	15.09%
	pred. 4	20.51%	5.40%	38.98%	20.53%	89.11%	59.76%	73.34%	41.03%	35.41%	30.98%	47.52%	36.51%
J48	pred. 3	74.57%	69.27%	81.84%	70.11%	83.39%	77.52%	97.89%	97.60%	81.36%	81.01%	81.00%	79.67%
	pred. 2	72.94%	64.92%	48.48%	39.50%	-	-	-	-	17.65%	6.67%	61.53%	61.53%
	pred. 5	-	-	3.17%	2.19%	46.16%	25.01%	-	-	36.36%	21.81%	0.00%	0.00%
	pred. 4	19.61%	4.35%	37.92%	16.28%	87.03%	82.61%	82.76%	78.57%	44.70%	38.85%	40.96%	18.19%
RIPPER	pred. 3	74.67%	72.04%	72.65%	71.48%	82.67%	71.64%	97.34%	96.12%	81.36%	80.43%	81.28%	79.36%
	pred. 2	-	-	56.41%	40.00%	-	-	-	-	12.90%	6.90%	53.33%	42.85%
	pred. 5	-	-	11.94%	6.06%	40.00%	18.19%	-	-	33.90%	25.00%	28.58%	28.58%
	pred. 4	9.75%	4.76%	35.79%	28.27%	82.50%	77.05%	69.57%	58.34%	46.02%	35.18%	36.43%	16.59%

Table III. Most extreme and least estimations of f-measures for pits activities with different ai methods

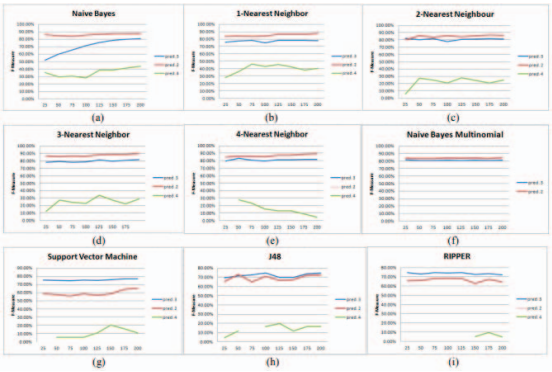
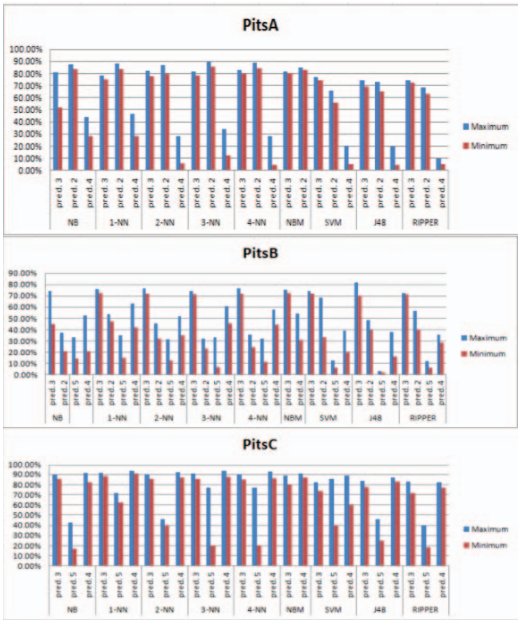


Fig. 2 The G- measure of various levels in different no. of terms



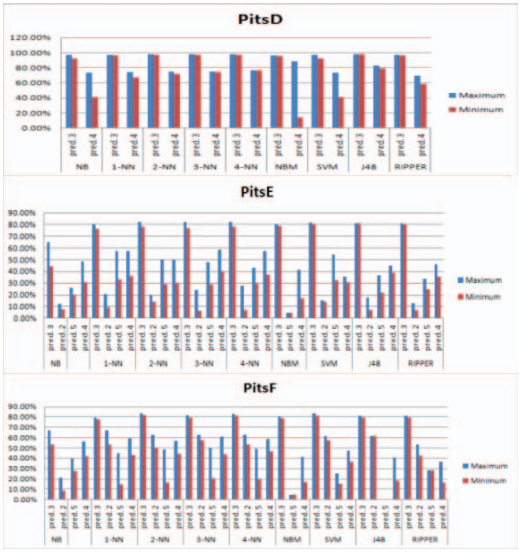


Fig 3. G-measures of predictions for pits project by NASA

VI. EVALUATION

As we prepared the classifier, we might want to gauge how precisely the classifier will detect future bug surveys. Right now, we talk about the various advances important to appropriately assess the classifiers.

a. *Training data and testing data*

Here we apply the method namely K-Fold Cross. This methodology first separates the assortment of the surveyed bug into disjoint preparing and testing sets, at that point prepares the classifier utilizing the preparation set. The result accuracy is given below,

Product	Name	*Non-severe* bugs	*Severe* bugs
Eclipse	SWT	696	3218
Eclipse	User Interface	1485	3351
JDT	User Interface	1470	1554
Eclipse	Debug	327	485
CDT	Debug	60	205
GEF	Draw2D	36	83
Evolution	Mailer	2537	7291
Evolution	Calendar	619	2661
GNOME	Panel	332	1297
Metacity	General	331	293
GStreamer	Core	93	352
Nautilus	CD-Burner	73	355

TableV Selected components result on severity of bugs

b. *Accuracy metrics*

There is four types of potential result. For example, a severe bug is anticipated effectively as a serious or mistakenly as a non-extreme bug. For the forecast of a non-serious bug, it is the other route around. We abridge these right and flawed expectations utilizing a solitary lattice, otherwise called the disarray framework, as introduced in table IV. This lattice gives a premise for the computation of numerous precision measurements.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

		Correct severity	
		non-severe	severe
Predicted severity	non-severe	tp: true positives	fp: false positives
	severe	fn: false negatives	tn: true negatives

Table VI. Bug severity of Confusion matrix

The general method for figuring the accuracy is by computing the level of bug surveys from the assessment set that are effectively grouped. Correspondingly, accuracy and review are generally utilized as assessment measures.

$$TPR = \frac{tp}{total\ number\ of\ positives} = \frac{tp}{tp + fn}$$

$$FPR = \frac{fp}{total\ number\ of\ negatives} = \frac{fp}{fp + tn}$$

The 2-D arrange framework for ROC and AUC curve probability is given below in the axis of 0.5 and 1.0. We can see that Classifier 1 exhibits irregular conduct. We additionally notice that Classifier 2 performs superior to arbitrary forecasts yet not comparable to Classifier 3.

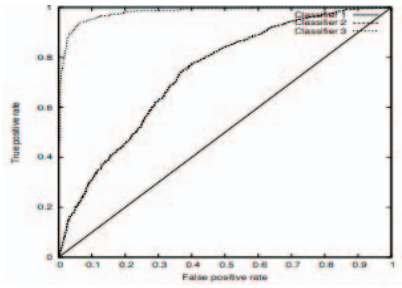


Fig. 3. Example of ROC curves

Along these lines, the zone underneath the curve is determined which fills in as a solitary number communicating the precision. On the off chance that the AUC is near 0.5 then the classifier is for all intents and purposes irregular, while a number near 1.0 implies that the classifier makes essentially flawless forecasts. This number permits increasingly sound conversations when looking at the precision of various classifiers.

VII. RESULTS AND DISCUSSION

Moreover, attributes that only find the most exact bug survey classified. At long last, we analyze what number of bug surveys we really require for preparing so as to acquire a classifier with sensible precision.

1. What classification algorithm would it be advisable for us to utilize when Anticipating bug survey severity?

Right now, I explored various algorithms: SVM, KNN, Naïve Bayes, Naïve Bayes Multinomial, J48 and RIPPER. Consequently, we look at the degrees of exactness of the algorithms.

In Fig 8 (a) and (b) shows that the curves of ROC of every algorithm indicating its exactness for separately an Obscuration and a Little person case. In Fig 3 shows that BM classified and VM classifier.

Product / Component	NB	NB Mult.	1-NN	SVM
Eclipse / SWT	0.74	**0.83**	0.62	*0.76*
JDT / UI	0.69	**0.75**	0.63	*0.71*
Eclipe / UI	0.70	**0.80**	0.61	*0.79*
Eclipse / Debug	0.72	**0.76**	0.67	*0.73*
GEF / Draw2D	**0.59**	*0.55*	0.51	0.48
CDT / Debug	0.68	**0.70**	0.52	*0.69*
Evolution / Mailer	0.84	**0.89**	0.73	*0.87*
Evolution / Calendar	*0.86*	**0.90**	0.78	*0.86*
GNOME / Panel	*0.89*	**0.90**	0.78	0.86
Metacity / General	*0.72*	**0.76**	0.69	0.71
GStreamer / Core	*0.74*	**0.76**	0.65	0.73
Nautilus / CDBurner	**0.93**	**0.93**	0.81	*0.91*

Table V ROC curve result

From these outcomes, we, in fact, notice that the Na¨ıve Bayes Multinomial classifier is the most exact in everything except GEF case. Right now, a number of accessible bug surveys for preparing are somewhat low and along these lines, we are managing an inadequately prepared classifier coming about normally in poor precision. Right now, The classifier dependent on SVM has a precision almost in the same class as the Na¨ıve Bayes Multinomial classifier.

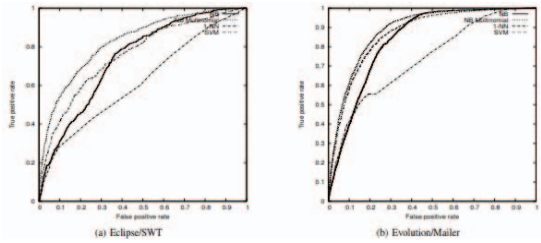


Fig. 4. Curves detector on existing algorithms

Product / Component	NB	NB Mult.	1-NN	SVM
Eclipse / UI	73	1	10	47
GNOME / Mailer	161	2	27	232

Table VI. EXECUTION OF COMPLETE EXPERIMENT

Table VI gives the complete experiment of the ML such as SVM, KNN, Naïve Bayes, Naïve Bayes Multinomial, J48 and RIPPER

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

The Naïve Bayes Multinomial classifier has the best accuracy (as measured by ROC) and is also the fastest when classifying the severity of reported bugs.

2. What number of bug surveys are important when preparing classifier based classifier expectations?

Other than the precision of the various classifier, the number of bug surveys in the arrangement set in like manner accept a critical activity in our process.

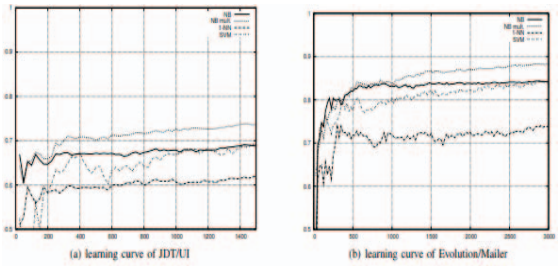


Fig. 5. Learning curves

By virtue of the JDT - UI section, for all ML classifiers curves measure adjust when the readiness set contains around 450 bug surveys (this suggests around 250 surveys of each reality) beside because of the classifier subject to Help Vector Machines. Comparable holds for the Headway – Mailer.

The Naïve Bayes and Naïve Bayes Multinomial classifiers are able to achieve stable accuracy the fastest. Also, we need about 250 bug reports of each severity when we train our classifier.

3. What would we be able to find from the subsequent classification algorithms?

The proper extract classifiers from the survey of bug in the preparation stage. The table VII shows that the nov severe and severe eclipse and evolution mailer.

Case	*Non-severe*	*Severe*
Eclipse JDT UI	quick, fix, dialog, type, gener, code, set, javadoc, wizard, mnemon, messag, prefer, import, method, manipul, button, warn, page, miss, wrong, extract, label, add, quickfix, pref, constant, enabl, icon, paramet, constructor	npe, java, file, package, open, junit, eclips, editor, folder, problem, project, cannot, delete, rename, error, view, search, fail, intern, broken, run, explore, cause, perspect, jdt, classpath, hang, resourc, save, crash
Evolution Mailer	message, not, button, change, dialog, display, doesnt, header, list, search, select, show, signature, text, unread, view, window, load, pad, ad, content, startup, subscribe, another, encrypt, import, press, print, sometimes	crash, evolut, mail, email, imap, click, evo, inbox, mailer, open, read, server, start, hang, bodi, cancel, onli, appli, junk, make, prefer, tree, user, automat, mode, sourc, sigsegv, warn, segment

Table VII Significant Terms of NB

At the point when we view Table VII, we notice that a few terms fit in with our desires: "crash", "hang", "npe" ,"fail" and the like are acceptable markers of serious bugs. This can be clarified by the birthplaces of serious bugs which are normally simpler to portray utilizing explicit terms.

Each component tends to have its own particular way of describing *severe* and *non-severe bugs*. Thus, terms which are good indicators of the severity are usually component-specific.

VIII. FUTURE RESEARCH AGENDA

In upcoming enhancement, the creators will chip away at the accompanying exploration plan:

1. This investigation will be done on additional ventures/parts of programming undertakings to approve the relevance of AI systems.
2. The exploration will likewise be completed utilizing bunch making in terms of a worldwide and nearby word reference which uses in decreasing the count of highlights.
3. Non-Global lexicons can be set up for every seriousness step by consolidating area explicit activities.

References

[1] Luiz Alberto , Ferreira Gomes, Ricardo da SilvaTorres, Mario LúcioCôrtes " Bug report severity level prediction in open source software: A survey and research opportunities" in Information and Software Technology Volume 115, November 2019, Pages 58-78

[2] J. M. Zhang, M. Harman, L. Ma and Y. Liu, "Machine Learning Testing: Survey, Landscapes and Horizons," in IEEE Transactions on Software Engineering.

[3] Xiaoxue Wu, Wei Zheng, Xiang Chen, FangWang, Dejun Mu "CVE-assisted large-scale security bug report dataset construction method" in Journal of Systems and Software Volume 160, February 2020, 110456

[4] Pinak Patel, Siddharth Mal, Yash Mhaske "A Survey Paper on Fraud Detection and Frequent Pattern Matching in Insurance claims using Data Mining Techniques" in International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06 Issue: 01 | Jan 2019

[5] M. N. Pushpalatha, M. Mrunalini "Predicting the Severity of Closed Source Bug Reports Using Ensemble Methods" in Smart Intelligent Computing and Applications pp 589-597 2018

[6] Asmita Yadav, Sandeep Kumar Singh "S urvey Based Classification of Bug Triage Approaches" in Aptikom journal on computer science and information technology 2019

[7] Krishna A Patel, Prof. Rohan C Prajapati "A Survey-Vulnerability Classification of Bug Reports using Multiple Machine Learning Approach" in COMPUSOFT, An international journal of

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

advanced computer technology, 5 (3), March - 2016 (Volume-V, Issue-III)
[8] Neeti Joshi "Survey of Rapid Software Testing using Machine Learning" in International Journal of Trend in Research and Development, Volume 3(5) 91-93 2016.
[9] M.Chalapathi Rao, Dr.P.Suryanarayana Babu "A Complete Survey of Research in Bug Complexity and Data Mining" in IJRAR January 2019, Volume 06, Issue 1 988-991.
[10] T. Zhang, J. Chen, H. Jiang, X. Luo and X. Xia, "Bug Report Enrichment with Application of Automated Fixer Recommendation," *2017 IEEE/ACM 25th International Conference on Program* Comprehension (ICPC), Buenos Aires, 2017, pp. 230-240.
[11] Tian-Lun Zhang, Rong Chen, Xi Yang & Hong-Yu Zhu "An uncertainty based incremental learning for identifying the severity of bug report" in International Journal of Machine Learning and Cybernetics volume 11, pages123–136(2020)
[12] Panagiota Galetsi, Korina Katsaliaki, Sameer Kumar "Big data analytics in health sector: Theoretical framework, techniques and prospects" in International Journal of Information Management Volume 50, February 2020, Pages 206-21
[13] Rudolf Ferenc, Dénes Bán, Tamás Grósz, Tibor Gyimóthy "Deep learning in static, metric-based bug prediction" in Elsevier https://www.sciencedirect.com/science/article/pii/S2590005620300060 (2020)
[14] Sajedul Talukder and Zahidur Talukder " A SURVEY ON MALWARE DETECTION AND ANALYSIS TOOLS" in International Journal of Network Security & Its Applications (IJNSA) Vol. 12, No.2, March 2020 37-57
[15]Ammara Zamir, HikmatUllah Khan, Tassawar I qbal, Nazish Yousaf, Farah Aslam, Almas Anjum, Maryam Hamdani " Phishing web site detection using diverse machine learning algorithms" in https://www.emerald.com/insight/content/doi/10.1108/EL-05-2019-0118/full/html 2 January 2020
[16] Anjali Goyal, Neetu Sardana "Analytical Study on Bug Triaging Practices" in https://www.igi-global.com/chapter/analytical-study-on-bug-triaging-practices/252107 2020.
[17] B. Zhou, I. Neamtiu and R. Gupta, "Experience report: How do bug characteristics differ across severity classes: A multi-platform study," 2015 IEEE 26th International Symposium on Software Reliability Engineering (ISSRE), Gaithersbury, MD, 2015, pp. 507-517.
[18] S. Guo, R. Chen, M. Wei, H. Li and Y. Liu, "Ensemble Data Reduction Techniques and Multi-RSMOTE via Fuzzy Integral for Bug Report Classification," in IEEE Access, vol. 6, pp. 45934-45950, 2018.

Bibliography



Master Muppala Sunny Chowdhary graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering has excelled in this stream. His research interest is Etymology, Android Applications in real life, Neural Networking, Data Analysis.



Ms R. Aishwarya, She is working as an Assistant Professor in the stream of Computer Science and Engineering at Sathyabama Institute of science and technology. She is an expert in Fog Computing and Cloud Computing. Her research interest is Cloud computing, Internet of Things. She is a Gold medallist in her PG in the year of 2016-2018.



Master Akula Abinay graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering has excellence in this stream. His research interest is Cyber Security, Android Applications in real life, Neural Networking, Data Analysis and Ethical Hacking.



Master Pasthapuram Harikrishna graduated from beloved University 'Sathyabama institute science and technology' in streaming of computer systems and engineering and has excellence in his stream. He has interest to do research in App development Machine learning,artificial neural network, IOT

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 05:54:14 UTC from IEEE Xplore. Restrictions apply.

Journal of Cluster Science (2020) 31:287–291
https://doi.org/10.1007/s10876-019-01626-4

BRIEF COMMUNICATION

Commercial Yeast Extracts Mediated Green Synthesis of Silver Chloride Nanoparticles and their Anti-mycobacterial Activity

Anbarasu Sivaraj[1] · Vanaja Kumar[1] · Revathy Sunder[1] · Krupakar Parthasarathy[1] · Govindaraju Kasivelu[2]

Received: 28 May 2019 / Published online: 6 July 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
Green synthesis of silver chloride nanoparticles (*AgCl NPs*) using commercial yeast extract has been carried out. The physico–chemical characterizations of *AgCl NPs* were carried out by UV–visible spectroscopy, XRD, FTIR, HR-SEM equipped with EDAX and TEM. In vitro efficacy of anti-mycobacterial properties of *AgCl NPs* were determined by agar well diffusion and Luciferase Reporter Phage (LRP) assay were used against *Mycobacterium smegmatis* (MC^2155) and *Mycobacterium tuberculosis* H37Rv respectively. This shows that the *AgCl NPs* have potential anti-myobacterial activity against *Mycobacterium tuberculosis* H37Rv.

Keywords Yeast extract · AgCl nanoparticles · LRP · Anti-mycobacterial activity

Introduction

Tuberculosis (TB) is a transmittable airborne disease caused by *Mycobacterium tuberculosis* (Mtb). As stated by WHO global TB report (2018), 10 million people (range 9.0–11.1 million) were estimated to have infected with TB in which around 558 000 patients developed Multiple Drug Resistant tuberculosis (MDR/RR-TB) and 1.3 million died from the disease around the world [1, 2]. The alarming increase in the rate of Multi Drug Resistant (MDR) and Extensively Drug Resistant (XDR) TB pose grave trouble around the world. Emergence of new drug candidate to fight against the drug resistant TB is thus necessary to design effective therapeutic strategies. Theranostic nanomedicine is emerging as a promising therapeutic paradigm for treating various diseases. Further, it takes more advantages of the high capacity of nanoplatforms to carry drug molecules for both imaging and therapeutic functions. The resulting nanosystems, able to disease diagnosis, drug delivery and monitoring of therapeutic function are expected to play an important role in the personalized nanomedicine, and much investigate effort has been devoted toward that goal [3]. Nanomedicine research has witnessed unprecedented growth in recent years. There is increasing optimism that nano-medicine, will bring significant advancement in the area of disease diagnosis and treatment [4]. Various biomedical applications include drug delivery, disease diagnosis, nutraceuticals, etc. [5–7]. For centuries, silver is known for its bactericidal property against a different group of bacteria and has been used for many years as an antimicrobial substance in the medical field [8, 9]. In recent years, resistance to antibacterial agents by pathogenic bacteria has become a major challenge for the health care industry. Research have endeavored to make use of unicellular and sub cellular supports as possible eco-friendly nanofactories for the synthesis of various metal and metal oxide nanomaterials [10]. In the present study, green preparation of silver chloride (*AgCl*) nanoparticles using commercial yeast and its anti-mycobacterial property against *M. tuberculosis* H37Rv and *M. smegmatis* has been investigated in detail.

Results and Discussion

MDR is a growing problem in the infectious diseases treatment and antibiotic resistance for various bacterial pathogens. In the field of nanoscience, allowing the

✉ Govindaraju Kasivelu
govindtu@gmail.com

[1] Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai 600 119, India

[2] Centre for Ocean Research, Sathyabama Institute of Science and Technology, Chennai 600 119, India

synthesis of nanomaterials that can be assembled atom by atom into complex structure for treatment of various diseases is carried out. Novel technologies are devoted to understanding the mechanisms of disease for the design of new drugs. Traditionally, silver was well known for its bactericidal property and for centuries it has been used for prevention and control of different infections, silver nanomaterials are attracting much interest because of their effective anti-microbial activity.

Commercial yeast extract (0.5 g) treated aqueous 1 mM concentration of $AgNO_3$ solutions were incubated overnight. After, the overnight incubation, the color changed into darkish brown which indicates the formation of *AgCl NPs* by visual observation (Fig. 1 inset). The absorption spectra of *AgCl NPs* formed in the reaction media showed strong absorbance peak at 423 nm (Fig. 1a). This may be due to the surface plasmon resonance excitation of the *AgCl NPs* [11]. Interaction of the biological material and metal salt concentration is known to influence nanoparticles synthesis and nanoparticles silver exhibit unique and tunable optical properties on account of their SPR, extremely sensitive to the nature, size distribution and shape of the nanoparticles [12].

The crystalline nature of the *AgCl NPs* was confirmed by powder XRD (Fig. 1b). The peaks observed at 2θ range 28.12°, 32.36°, 46.34°, 54.92°, 57.54°, and 76.69° which corresponds to the (111) (200) (220) (311) (222) and (420) planes of the cubic silver chloride nanoparticles (ICDD file no. 00-001-1013) [13]. Crystallite size estimation was made from the line broadening spectrum of the reflection using the Debye Scherer formula. The calculated crystallite size of the *AgCl NPs* was around 17 nm.

FTIR spectroscopy results of both yeast extract and *AgCl NPs* were shown in Fig. 1c. For yeast extract showed

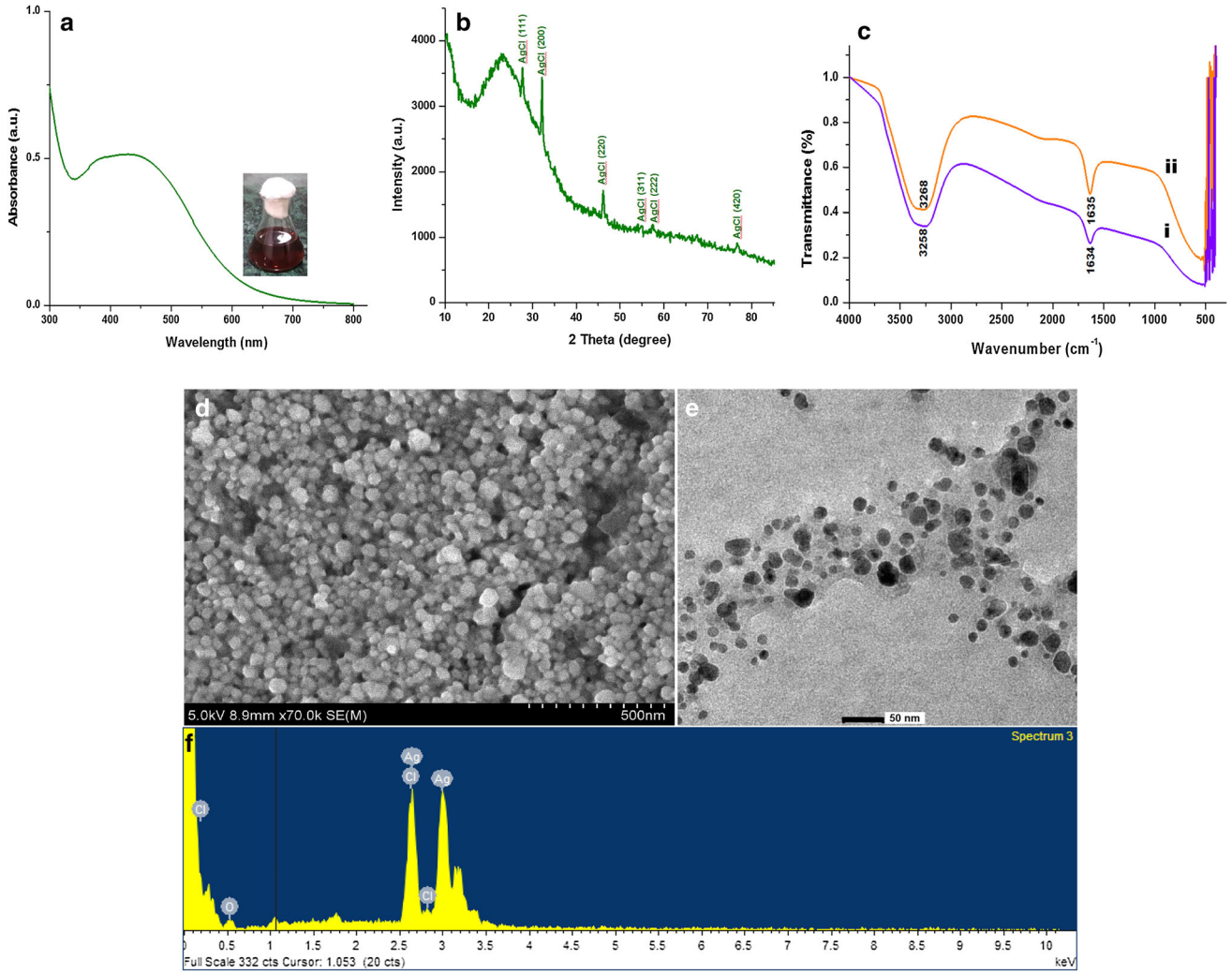


Fig.1 **a** UV–vis spectra of *AgCl NPs*. Inset shows colloidal *AgCl NPs*; **b** XRD spectrum of *AgCl NPs* **c** FTIR spectrum of commercial yeast (purple line) and *AgCl NPs* (orange line); **d** HR-SEM image of *AgCl NPs*; **e** TEM image of *AgCl NPs*; **f** EDAX spectrum of *AgCl NPs* (Color figure online)

the peaks were observed at 3258 cm^{-1} (N–H stretch, primary amine) and 1634 cm^{-1} (amide C=O stretch) [Fig. 1c(i)]. For *AgCl NPs* were observed at 3266 cm^{-1} (N–H stretch, primary amine) and 1635 cm^{-1} (amide C=O stretch) [Fig. 1c(ii)]. The comparison of the both resultant spectrum between the yeast extract and the *AgCl NPs* slight shift change was observed in the peak position as well as the absorption bands and no shift change was observed at 1635 cm^{-1} [14]. Primary amine functional group present in the yeast extract protein have a stronger affinity to bind with metal ions surface, facilitates reduction of silver ion into *AgCl NPs*.

Colloidal *AgCl NPs* were observed using HR-SEM which clearly shows that the spherical structure of size ranging from 9 to 51 nm (Fig. 1d) further, the elemental composition of the *AgCl NPs* was obtained from EDAX measurement (Fig. 1f). Two major peaks with high intensity signals of Ag and Cl were observed, which indicated the presence of *AgCl NPs*. TEM images confirmed that the *AgCl NPs* were spherical in shape and the size of the particles was ca. 17 nm (Fig. 1e).

Nanoparticles have the ability to destroy bacteria by penetrating the cell membrane, followed by interfering in intracellular ATP production and inhibition of DNA replication. It has been reported to induce oxidative stress in bacteria [15]. Few studies showed that chemical and microbial Ag NPs having mycobactericidal activity [16, 17]. In this study, yeast extract tested against *M. smegmatis* Mc^2155 there is no inhibitory activity (Fig. 1g), whereas, *AgCl NPs* were tested for their antimycobacterial activity against *M. smegmatis* Mc^2155 and the results showed that *AgCl NPs* inhibited around 13 mm diameter at 37 µg/mL concentrations (Fig. 2a). Silver nanoparticles were synthesized using *Psidium guajava* L showed activity against *M. smegmatis* (12.53 ± 0.42 mm) [18]. In the present work, *AgCl NPs* have show 92% reduction at concentration of 37 µg/mL and standard drug rifampicin (2 µg/mL concentration) showed 95% reduction against *M. tuberculosis* H37Rv (Fig. 2b). Banu and Rathod have reported that biologically synthesized silver nanoparticles using *R. stolonifer* have showed minimum inhibition concentration of 12.5 µg/mL against resistant strains of *M. tuberculosis* [19]. Agarwal et al. [20] synthesized silver nanoparticles using *Cucumis sativus* plant extract which showed inhibition against clinical isolates of *M. tuberculosis* within the MIC range of 7.8 to 15.6 µg/mL. AgNPs synthesized using *B. prionitis* leaf extract and *S. cumini* bark extract and on exposure, these nanoparticles were observed to inhibit 60–85% of dormant mycobacteria [21]. Biosynthesized silver nanoparticles using *L. acidissima* plant extract have shown the inhibition against *M. tuberculosis* at 3.125 µg/mL [22]. Many studies have been reported for anti-tubercular activities of silver nanoparticles synthesized using plant extracts and few studies using microbial source such as bacteria and fungi. However, to the best of our knowledge this is the first report pertaining to *AgCl NPs* synthesized using yeast extract which possess anti-mycobacterial properties. Further, the molecular mechanism for the action of the nanoparticles and the tuberculosis strains are currently under investigation.

The present study shows that the commercially available yeast extract powder can be used as an effective reducing agent for the synthesis of *AgCl NPs* from silver nitrate solution. In this process, nanoparticles were formed in overnight incubation and this process is very inexpensive and eco-friendly method. In conclusion, the anti-mycobacterial property of biogenic *AgCl NPs* exhibits profound efficiency to inhibit both *M. tuberculosis* and *M. smegmatis* in vitro. The present study reveals *AgCl NPs* can act as novel anti-mycobacterial agent.

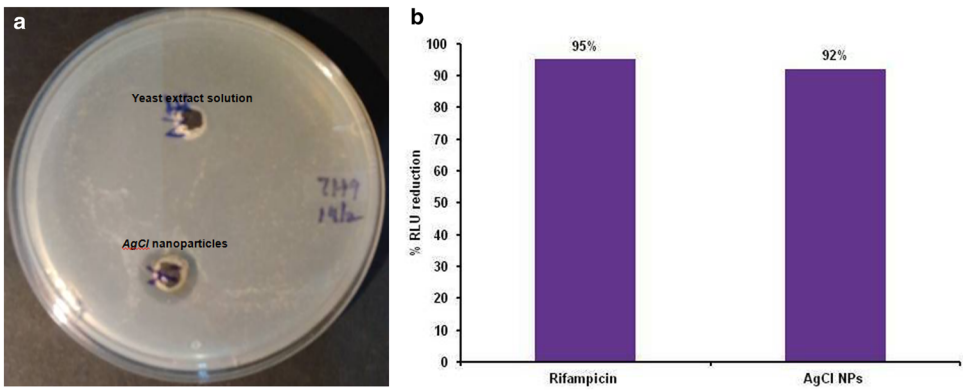


Fig.2 **a** Zone of inhibitory activity of commercial yeast and *AgCl NPs* against *M. smegmatis* and **b** RLU reduction of *AgCl NPs*

Experimental Procedure

Materials

Commercial yeast extract was purchased from Himedia, India. Silver nitrate ($AgNO_3$) was obtained from Merck. Test organisms, *M. tuberculosis* H37Rv, *M. smegmatis* MC^2155 and Mycobacteriophage (phAE 202) were kindly provided by Dr. Vanaja Kumar, Scientist G, Department of Bacteriology, National Institute for Research in Tuberculosis (NIRT), Indian Council of Medical Research (ICMR), Government of India, Chennai, India.

Synthesis of AgCl NPs Using Commercial Yeast

Yeast extract (0.5 g) were dissolved in 100 mL of deionized water and sterilized at 121 °C for 15 min. Freshly prepared $AgNO_3$ solution (1 mM) was added into sterilized yeast extract solution at 90:10 mL ratio respectively and then incubated overnight at room temperature. After incubation, the solution color changed to dark brown, aliquots of the reaction solution were measured using a UV–vis spectrophotometer operated at a resolution of 1 nm. Colloidal *AgCl NPs* solution were filtered using sterilized syringe filter (0.45 µ PES filter, Himedia) and then stored at 4 °C until testing for anti-mycobacterial activity.

Physico–Chemical Characterization of AgCl NPs

The formation of *AgCl NPs* observed periodically by UV–vis spectral measuring in the range of 300–800 nm (Shimadzu UV-1800, Japan). In order to find out the crystalline nature of metallic nanoparticles, sample were subjected to XRD analysis (BRUKER AXS, D8 Discover) operating in transmission mode at 40 kV and 30 mA with Cu Kα radiation. FTIR performed to find out the possible bio molecules responsible for the formation of nanoparticles. The samples (yeast extract and *AgCl NPs*) were mixed with KBr to make a pellet and subjected for FTIR spectral analysis (BRUKER IFS 66 V) at the range of 500–4000 cm^{-1} and the resolution was kept as 4 cm^{-1}. Surface morphological and elemental composition analysis of the biosynthesized *AgCl NPs* were carried out using HR-SEM equipped with EDAX (Hitachi-S4800). Analysis of size and shape of the *AgCl NPs* were carried using TEM. A small volume of *AgCl NPs* on carbon coated copper grid, air dried and measured on the FEI Tecnai Sprit G2 instrument with an accelerating voltage of HT 120 KVA.

Preliminary Screening for Anti-mycobacterial Activity

A suspension of *M. smegmatis* was prepared by dissolving a loopful of 24 h culture into a bijou bottle containing 0.3 mL of sterile Middlebrook 7H9 broth with glass beads. The mixture was vortexed for 30 s, and then the suspension was made upto 1 mL with sterile Middlebrook 7H9 Broth. The suspension was slightly shaken and kept undisturbed for 10 min to allow the clumps to settle down. 200 µL of *M. smegmatis* suspension was added with 5 mL of molten soft agar (1% agar) at 45 °C and poured immediately onto fresh Middlebrook 7H9 agar base plate. After allowing the agar to solidify, well was cut into the plate and 100 µL of *AgCl NPs* were added into well. Then the plates were sealed and incubated at 37 °C. The zone of inhibition was measured as the diameter of clear area without bacterial lawn around the edges of the wells.

Anti-mycobacterial Activity

Luciferase reporter phage (LRP) assay method is used for high throughput screening of compounds/nanoparticles for their anti-mycobacterial activity. This method is rapid, inexpensive and less laborious whereas BATEC methodology is costly, cumbersome and uses radioactive reagents [23]. *AgCl NPs* is considered as an anti-tubercular agent if 50% reduction in relative lights units (RLU) is observed when compared to the control using luminometer [24]. Briefly, About 400 µL of sterile middlebrook 7H9 broth (Himedia) and prepared Rif-7H9-2 µg/mL was transferred to control cryovials (C) and drug control cryovial (DC) respectively. 350 µL of sterile middlebrook 7H9 broth and 50 µL of *AgCl NPs* were added to test cryovial (TC). All cryovials were inoculated with 100 µL of *M. tuberculosis* H37Rv cell suspension (McFarland No. 2) and incubated at 37 °C for 72 h. After incubation, 50 µL of mycobacteriophage (phAE202) and 40 µL of 0.1 M $CaCl_2$ were added into all the cryovials and incubated for 4 h at 37 °C. Then, 100 µL of cell-phage mixture was transferred to a luminometer cuvette and added with 100 µL of D-luciferin substrate (Cayman chemicals, USA). The relative light unit (RLU) was measured immediately at 10 s integration time in luminometer (Lumat 9508, Berthold, Germany). *AgCl NPs* showed more than 50% reduction in RLU were considered as active against *M. tuberculosis* H37Rv. The percentage of reduction was calculated by using below mentioned formula:

$$\text{RLU reduction } (\%) = \frac{\text{Control RLU} - \text{Sample RLU}}{\text{Control RLU}} \times 100$$

Acknowledgements We thank the management of Sathyabama Institute of Science and Technology, Chennai for its strong support in research activities.

References

1. C. Jacob, S. Suvanand, S.S. Kuldeep, D. Lucica, B. Draurio, M. Andrei, M. Chen, P. Yogan (2014). *Bull. World Health Organ.* **92**, 459.
2. WHO, Global tuberculosis report (2018), https://www.who.int/tb/publications/global_report/en/
3. J. Xie, S. Lee, X. Chen (2010). *Adv. Drug Deliv. Rev.* **62**, 1064.
4. W.H. De Jong, P.J.A. Borm (2008). *Int. J. Nanomed.* **3**, 133.
5. R. Ducan (2003). *Nat. Rev. Drug Discov.* **2**, 347.
6. W.H. De Jong, R.E. Geertsma and B. Roszek, *Nanotechnology in medical applications*. Possible risks for human health, Report 265001002/2005, National Institute for Public Health and the Environment, Bilthoven, The Netherlands.
7. European Technology Platform on Nanomedicine, *Vision paper and basis for a strategic research agenda for nanomedicine.* (European Commission, Luxembourg, 2005). ISBN 92-894-6599-5
8. J.J. Castellano, S.M. Shafii, F. Ko, G. Donate, T.E. Wright, R.J. Mannari, W.G. Payne, D.J. Smith, M.C. Robson (2007). *Int. Wound J.* **4**, 114.
9. J.W. Richard, B.A. Spencer, L.F. McCoy, E. Carina, J. Washington, P. Edgar (2002). *J. Burns Surg. Wound Care* **1**, 11.
10. C.H. Luo, V. Shanmugam, C.S. Yeh (2015). *NPG Asia Mater.* **7**, 1.
11. P. Mulvaney (1996). *Langmuir* **12**, 788.
12. N. Ahmad, S. Sharma, M.K. Alam, V.N. Singh, S.F. Shamsi, B.R. Mehta, A. Fatma (2010). *Colloids Surf. B: Biointerf.* **81**, 81.
13. V. Gopinath, S. Priyadarshini, N. Meera Priyadharsshini, K. Pandian, P. Velusamy (2013). *Mater. Lett.* **91**, 227.
14. V. Gopinath, D. Mubarak Ali, S. Priyadarshini, N.M. Priyadharsshini, N. Thajuddin, P. Velusamy (2012). *Colloids Surf. B.* **96**, 69.
15. L. Wang, C. Hu, L. Shao (2017). *Int. J. Nanomedicine* **12**, 1227.
16. S. Mohanty, P. Jena, R. Mehta, R. Pati, B. Banerjee, S. Patil, A. Sonawane (2013). *Antimicrob. Agents Chemother.* **57**, 3688.
17. R. Singh, L.U. Nawale, M. Arkile, U.U. Shedbalkar, S.A. Wadhwani, D. Sarkar, B.A. Chopade (2015). *Int. J. Antimicrob. Agents* **46**, 183.
18. R.K. Jivan, M.M. Ramjan, S.K. Ambadas, M.S. Bhimrao (2014). *J. Microbiol. Biotechnol. Res.* **4**, 14.
19. A. Banu, V. Rathod (2013). *J. Nanomed. Biotherapeut. Discov.* **3**, 110.
20. P. Agarwal, A. Mehta, S. Kachhwaha, S.L. Kothari (2013). *Adv. Sci. Eng. Med.* **5**, 709.
21. R. Singh, L. Nawale, M. Aarkile, S. Wadhwani, U. Shedbalkar, S. Chopade, D. Sarkar, B.A. Chopade (2016). *Int. J. Nanomed.* **11**, 1889.
22. N.P. Bheemanagouda, T.C. Taranath (2016). *Int. J. Pharm. Bio. Sci.* **7**, 89.
23. Papitha N,Jayshree N, PrabuSeenivasan S,Kumar V (2013). *Int. J. Pharm. Sci. Rev. Res.* **20**, 135.
24. National Workshop on LRP 2011 Department of Bacteriology, National Institute for Research in Tuberculosis (NIRT) Chennai India

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cognizance the Action of Human by Applying Enhanced Techniques

R. Aishwarya[*]
Assistant Professor
aishwaryarajkumar7@gmail.com

Muppala Sunny Chowdhary[2]
Student
chowdarysunny89@outlook.com

Akula Abinay[3]
Student
akulaabinay@gmail.com

Kethamukkala Phanindhra[4]
Student
phanindra1613@outlook.com

Nadupuru Yamini[5]
Student
yamininadupuru2510@gmail.com

Reema Rose Toppo[6]
Student
reema.rose@outlook.com

[*,2,3,4,5,6] Sathyabama Institute of Science and Technology

Abstract: **-Human conduct start`s daily practice by commerce with others. Commerce in the sense that it has much social companionship mainly through mobile phones. So far, we gather the human emotion through sentiment analysis by using a hybrid approach to determine the emotion through the text-based operation. Using EEG Signal, we can identify the voice-based data with the help of phone call log in k-means clustering method to gather the voice data. Using this data set, snowball sampling is used to locate the hidden information and describing the emotions of humans. In this paper, the analysis of emotion is done using three techniques such as facial recognition, speech recognition, textual analysis. For the facial recognition by predicting human emotion through webcam. The voice recognition of the human emotion is determined by recording the human voice and comparing with the frequency where the frequency has been collected and stored in the data-sets. The textual analysis is based on the classification of the statement in positive, negative and by general identification. This helps to identify whether the statement is activate or deactivate. From this above classifications we can determine the emotions of the human. The performance can be proven with our proposed work to find different emotions of humans by using existing enhanced techniques.**

Keywords-Enhanced techniques, Facial recognition, Speech recognition, Textual analysis, Emotion spacing, HMM (Hidden Markov Model), SVM(Support vector machine)

I. INTRODUCTION

The human emotion analysis is a fundamental part of today's affective computing world. "Affect" mean emotion and "computing" mean calculating or measure of the emotion. Calculation of the human emotion is nothing but making or designing the system or device that recognize, process, stimulate and interrupt the emotion of the human. We can make human and the machine interacts for a better understanding of human emotion. The data can be of various forms like facial expression, voice frequency, linguistics text shows in table 1.

	Positive	Negative
Activate	Joyful Happy Playful	Anxious Angry Tense
Deactivate	Relaxed Tranquil Serene	Sad Ashamed Grieving

Table.1. Different type of Human Emotions

In order of classifying the human emotion, there are divided into four categories which are (positive, activated), (negative, activated), (positive, deactivated), (negative, deactivated). In positive, activated there are happy, joyful, playful etc. In the negative, activated there are angry, tense, anxious etc. In deactivated, positive there are relaxed, serene, tranquil, etc. In deactivated, negative sad, ashamed, grieving, etc, the above classification is based on the terms positive, negative, activated and deactivated. In classifying the human emotions, we are using three enhanced techniques which are

a. Facial reorganization
b. Speech reorganization
c. Textual analysis

II. SENTIMENTAL ANALYSIS

The sentiment analysis contains five steps. They are:

• **Data collection**: the First step is to collect data from users who involved in using mobile phones

 ISBN 978-1-7281-6202-7

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

for both text-based and voice-based in the personal and social network.

• **Text preparation**: Non-textual contents will be eliminated ex- emoji, symbol etc.., to get complete text-based information.

• **Sentiment detection**: The information extract from the reviews are examined. Sentence with subjective
expressions are to be absorbed and sentence with objective expression are to be neglected.

• **Sentiment Classification**: The classification based on the positive, negative, good, bad, agree, disagree and it can be made by multiple points.

• **Output**: The main goal is to transfer the unstructured text into meaningful information. The tools we are using here is Linguistic Inquiry and Word Count, it used to analyse not only positive, negative but also analyse emotional, structural and cognitive components of text based on the use of dictionary containing words shows in Table 2.

LIWC Dimension	Your Data	Personal Texts	Formal Texts
Self-citation (I,me,my)	0.74	11.4	4.2
Communal words	5.19	9.5	8.0
Optimistic words	0.74	2.7	2.6
Gloomy emotions	0.74	2.6	1.6
Overall insight words	5.93	7.8	5.4
Articles(a,an,the)	4.44	5.0	7.2
Big words(>6 letters)	39.26	13.1	19.6

Table.2.Example for sentimental analysis using LIWC

III. ENHANCED TECHNIQUES

In Table 3. Classification of Human Emotion is not an easy task, where different types of methods and algorithms are proposed to retrieve the final specification of the human emotion. In the following below techniques,

a. FACIAL RECOGNITION
b. SPEECH RECOGNITION
c. TEXTUAL ANALYSIS

Sentiment Classification Approach	Features	Applications
Machine Learning	Speech information, Negation and Opinion	Ability to adapt and create a trained module
Lexical Based	Manual Construction and Dictionary Based	Wider Coverage
Hybrid Based	Sentimental Word i.e., Combination both Machine learning and Lexical based	Sentiment Detection and Measurement at the Concept Level

Table.3. Approaches of sentimental analysis

A. FACIAL RECOGNITION

The first enhanced technique which is in the discussion is facial recognition where we identify the emotion of the human on the live webcam approach. In the process of classifying the emotion of the human, we need to follow the process for arriving at the final output shows in Fig.1.

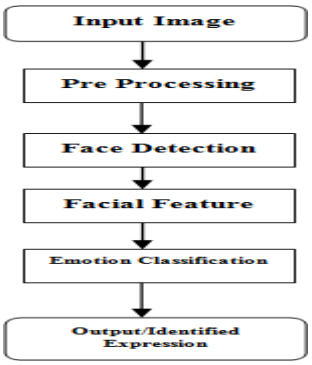


Fig.1**.** Extraction of Human Emotion by facial recognition

As mentioned in the diagram the steps to be followed for achieving the final result which is the actual emotion of the human are

a. Input images and pre-processing the given image
b. Expression compound
c. Network structure
d. Emotion analysis
e. Evolution of emotion space
f. Emotion recognition from the image

Input-Output image Pre-treating

As part of the discussion, the concerned input-output of the image pre-treating is a bright intense image where the pre-treating improves the image data and suppresses the unwanted data advances of some image features for further treating.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

Expression compound

The key piece of the demeanour union is the term demeanour framework, where the Facial picture mix is made out of disfiguring the wire plot sample between outward appearance frameworks and depicting the surface in every polygon of the primary picture onto the outside of the distorted wire edge model. The twisting standards are characterized to re-sanction facial strong exercises. Ekman and Friesen crumbled the facial solid exercises into 44 principal exercises called "Activity Sections" (AUs) and delineated the outward appearance as mixes of AUs. Occasions of Action Sections show up in the beneath table Fundamental verbalization can be portrayed into 6 groupings like Surprise, Happiness, Anger, Fear, Sadness and Disgust. Each feeling in like manner can be recognized by the mix of Action Sections. Getting ready of the framework relies upon these 6 sorts of feeling states appears in Table 4.

AU No.	FACS Name	AU No.	FACS Name
AU1	Inner brow raiser	AU14	Dimpler
AU2	Outer brow raiser	AU15	Lip corner depressor
AU4	Brow lower	AU16	Lower lip depressor
AU5	Upper lid raiser	AU17	Chin raiser
AU6	Cheek raiser	AU20	Lip strecher
AU7	Lid tightener	AU23	Lip tightener
AU9	Nose wrinkler	AU25	Lip part
AU10	Upper lid raiser	AU26	Jaw drops
AU12	Lid corner puller		

Table.4. Action sections for Expression

Network Structure

As appeared in Fig 2 underneath the frame structure which is the pre distortion sort with 5 stages. The main stage and fifth stage have a comparative number of sections. The covered stages in the second and fourth stages have 19 sections to restrict the mix goof esteem. The focusing stage (third stage) has 2 sections for this circumstance. This framework comprehends a character depicting and each section of information and yield stage thinks about the Action Section. The section number strategies AU number and section worth is AU power. Feeling space can be created in the module stage as the 2-D space by back propagation. Hypothesis effect of the neural framework can be used as the expansion of facial verbalization. There are 44 Action Sections in FACS, yet we picked 17 sections which give strong effect on enthusiastic enunciation. These frameworks have quite recently appeared in the underneath table. AU frameworks are institutionalized to worth and given to the neural system. The section capacity of first, third and fifth stages is a straight capacity to keep the wide extent of framework space. That of the second and fourth stages is the sigmoid capacity for non-straight depicting. The initial three stages from commitment to third can comprehend feeling affirmation system from enunciation frameworks. Feeling state is picked as the (x,y) point in emotion space in the middle stage. The last 3 stages can comprehend enunciation association system from feeling state. Face explanation picture is consolidated by the AU frameworks picked by the depicting from the (x,y) regard in feeling space.

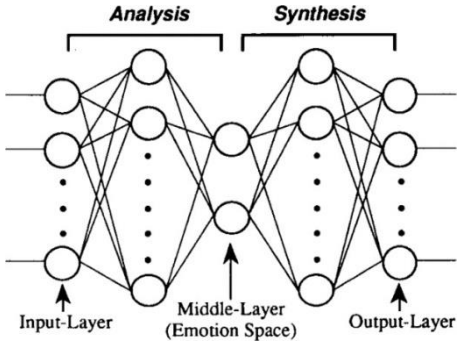


Fig.2. Triple-stageed Neural Network Structure

Emotion Analysis

In Fig.3 . 2-D feeling space delivered in the middle stage by back propagation has showed up in Figure 4. Each line suggests the locus line in the middle stage when the AU framework set of each enunciation in the data stage changes from unprejudiced to its most prominent power. Each spot is arranged in each 10%. You can see the dispersal of each feeling is uniform and fair is arranged in the center position. Ecstasy, Sadness, Surprise and Anger are arranged despite what might be expected situation in the inclination space separately. Exactly when any self-confident explanation which is portrayed with a blend of Action Sections is given to the data stage, the zone of this verbalization in the inclination space is picked by the estimations of focus stage of the neural framework. What's more, the development of verbalization is depicted with the point improvement in the inclination space. The territory of each feeling is reasonable and like a couple of feeling models proposed by an analyst. By non-direct depicting, the ramifications of each rotate are not uniform.

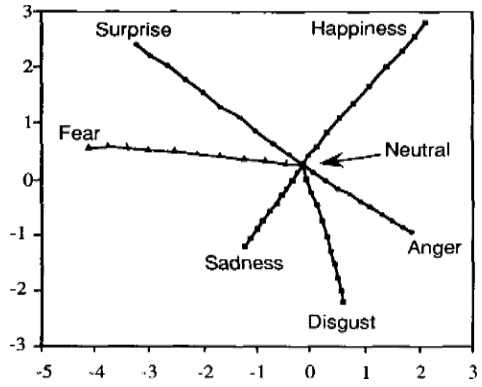


Fig. 3. 2-D Feeling Space

ISBN 978-1-7281-6202-7

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

Evolution of emotion space and recognition
Mental prior of feeling space is carryout by providing the face image to themes which are harvest by the point in feeling space. Test pictures for estimation are delivered by the concentrations in feeling space which arranges in the focal points of square squares. The quantity of squares is 81(9 by 9). The figure speaks to how the test picture is made. From the beginning, a square district in feeling space which consolidates most of the basic inclination positions is segregated into 9 by 9 square systems. facilitate (x,y) in the point of convergence of each square is given to the inside stage of the neural framework and from that point onward, the Action Section frameworks appear in the yield stage shows in Fig.4.

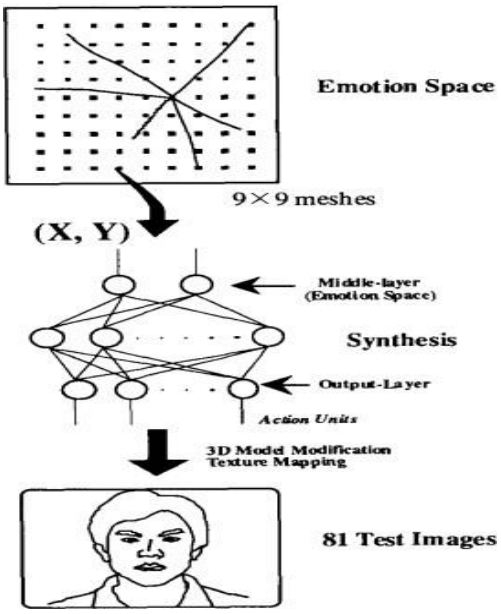


Fig.4. Subjective Test Image compound through Emotion Space

Framework depicting from AU framework to emotion space and that pivot depicting can be recognized by the 5-stageed neural framework. For the human-machine instinctive system with a face picture, appearance frameworks should be picked by the isolated component centers from the real picture. A ground-breaking computation to see feature motivations behind the face isn't found now, yet there is tip top estimation in case of fixed lighting state using edge, concealing and development information. In Fig.5 shows the segment centers for AU affirmation. Focuses that are arranged on the base of ears and top of the nose are used to pick the head development. focuses are the part centers to pick the facial appearance. These centers are recognized by the self-loader affirmation procedure, so manual change by the head is possible if significant. Exactly when Ai is the AU feature vector and W is its weighting vector, the entirety of the explanation can be portrayed by vector Dd shows in Fig. 5.

$$Dd=A^* W= (A1, A2, An) * W$$

Exactly when the detachment of feature centers from an unbiased situation in the image is Dp W is picked to restrain the estimation of Error=|Dp-Dd|. AU framework power is delineated by the segment of vector W. There are associations between AU frameworks Ai so extraction is acted in the solicitation for mouth, nose and eyebrow to restrain the calculation time. The incorporated picture from normally isolated AU framework works out to be for all intents and purposes proportional to the impression of the main picture by test and abstract test.

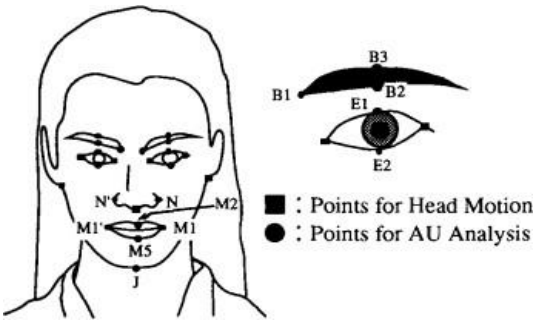


Fig.5. Feature Points for Expression Recognition

B. SPEECH RECOGNITION

Here, we have used the module/data-set namely known (speech recognition). The executable code helped in converting the audio file or live speech into a corresponding text file. the algorithms which are used for working of speech recognition through language modelling and acoustic modelling. Language modelling will correlate with the sound word sequences to distinguish the words which sound different and similar acoustic modelling makes up the relation between the audio signal and linguistic sections of speech.

The recognition of the speech is taken in the wave-form where we split up the wave signal utterances by silence and then we recognize what has been said/uttered in the speech signal with each utterance. For getting the audio file/speech signal in the text form we follow the below-mentioned process:

a. Speech input
b. Feature extraction
c. Feature selection
d. Classifier
e. Reorganized output

Speech input:
The input is of the waveform where we convert the speech signal into the text file with the help of the proposed algorithm and executable code.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

Feature extraction and selection:
In the concept of the matching process, there is a very large number of frameworks where we optimized the frameworks into less frame where the numbers are calculated from analyzing the speech by dividing into frames. For each frame the typical execution part takes about 10-milliseconds length. The feature selection is based on the model. A model is described on the bases of the mathematical object which gathers the common or similar attributes of the words which are spoken or given in the audio file /speech recognition.

Classifier:
In speech recognition we used three types of classifiers namely

a. HMM
b. SVM

C. TEXTUAL RECOGNITION

Textual analysis/recognition is another enhanced technique where we classify the emotion of the human where there are many approaches for getting the final output. Here are some of the approaches for the textual analysis/recognition. The text-based approach by computational methods for identifying human emotion shows in Fig 6.



Fig.6. Textual analysis of the human emotion

1. **Lexicon based methodology:** The lexicon-based methodologies exploitation lexical choices are separated into 2 subtypes: word dictionary-based and corpus-based. The lexicon essentially based methodologies start with some predefined lexicon of passionate words thus utilize shifted measures like term recurrence, word count(called insights) or word synonyms(called semantics) and so forth to name the sentences inside the information. Most applied arithmetic methodologies utilize Latent semantics Analysis (LSA) and even their variety is used for breaking down the association between a gathering of archives and in this way the terms in these reports in order to supply substantive examples related with records and terms. The dictionary approach fuses a low worth anyway at consistent time substantiating a lexicon is extreme. The corpus basically based methodology uses any broad information for feeling examination. Here the corpus (information) is first clarified by adhering to a gathering of unique guidelines, from a book, for overseeing a semantic correspondence feeling examination. Watchword basically based methodologies plot a gathering of predefined terms to order the content into enthusiastic classifications. Burly used Word Net-Affect conjointly for checking feeling words inside the features. mysticism essentially based methodology utilizes the association between the terms and thusly the EmotiNet and model situation as a succession of activities and their comparing enthusiastic affect. This methodology is also trailed by fine-grained feeling discovery.

2. **Machine learning approach**: The AI approach relies upon AI calculations which will gain from data by making utilization of etymological alternatives of content. They're additionally isolated into following:

A. Supervised Learning

B. Unsupervised Learning

C. Semi-Supervised Learning

a. **Supervised Learning**: These calculations type a perform (model) bolstered the PC document and exploited this capacity to take decisions of the best approach to delineate longer-term information to adequate yield. The SVM could be a conventional methodology during this respect. Not many scientists have hostage on the far side these antiquated ways to deal with extra affordable and dependable ways like CRF. The regulated learning models are any of the accompanying; call tree classifiers: utilize the progressive algorithmic decay of instructing data bolstered the bases of the value of the properties; till some leaf hubs are arrived at that contain values for the grouping reason; Rule basically based classifier: here the order is formed on the reason of some arrangement of rules. The states in superlative customary sort on the choices are depicted by the manners angle while the classification names are spoken to by paw side of the standard; Linear Classifiers: it arranges the feelings by making a decision bolstered the value of a direct blend of qualities of the information content. These qualities are alluded to as highlight esteems and are depicted in a surpassing vector type known as highlight vectors. It comprises of the many anticipated models, among them is that the help vector machines and Neural Networks; Probabilistic Classifier: expect each classification to be a piece of the blend that gives the opportsectiony of inspecting a chose term for that segment. It any comprises of following classifiers; Naive Bayes classifier, that figures back possibility bolstered the circulation of words in a surpassing archive: Bayesian system, is A non-

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

cyclic chart whose hubs speak to stochastic variable and edges speak to restrictive states; most entropy, utilizes cryptography to change over marked capabilities to vectors. This vector at that point figures loads for each component that is consolidated to work out the classification for each list of capabilities shows an examination between progressive versus level order of emotions.

b. **Unsupervised Learning**: these calculations endeavour to understand the concealed structures inside the info document and abuse those structures map the unlabeled information to feeling classifications.

c. **Semi-Supervised Learning**: There are a few works grew any place marking is finished precisely through hash tags and so forth. Semi-administered calculations utilize this idea of programmed marking and follow the consequent 2 methodologies: Bootstrapping and far off administration.

For executing the code we need the accompanying modules which are fundamental for the usage of the code

1. Pandas
2. Scipy
3. Scikit-learn
4. Beautifulsoup4
5. NLTK
6. Cython
7. Gensim

IV. RELATED WORK/METHODOLOGY

For classifying the emotion of the human we have adopted three enhanced techniques where they are

a. Facial recognition
b. Speech recognition
c. Textual analysis

The main aim/motto of this proposed work is to achieve the real-time emotion of the human for the existing enhanced techniques. The way we related the three enhanced technique is that from the facial recognition we classify the emotion of the human which is a partial output of our classification. From the other enhanced technique which is speech recognition, we classify the emotion and we obtain the textual form of the speech and from the proposed algorithm we find whether the text is positive or negative and from the general identification of the text is activated or deactivated And we finally arrive at another partial output of the human emotion. By comparing the two partial output which we obtained from executing the algorithm/source code we arrive at the final output which we generalize as the emotion of the human where it is the real-time emotions of the human.

V. VISUAL SURVEY/TESTING THE PROPOSED PROGRAM

We tested the algorithm and the source code where the algorithm of each enhances technique is stated below:

A. FACIAL RECOGNITION

```
PROCEDURE Face recognition
INPUT: Size, Cascade Classifier, Video Capture,
test.jpg
OUTPUT: Faces, Key, Variables.
VARIABLES: Classifiers, Webcam, sub_face
BEING
Significant cv2 and lable_image
WHILE TRUE
DO
   Read webcam
   Resize the image to speed up the detection
   mini=cv2 resize(im,(int(im.shape[1]/size),
int(im.shape[0]/size)))
   Detect faces
FOR f==faces
   Scale the shape size backup using
            (x, y, w, h) = [v * size for v in f]
            sub_face = im[y:y+h, x:x+w]
            Write (FaceFileName, Sub_face)
            Result form the label_imagefile,i.e,
classification Result
END FOR
 cv2.imshow('Capture',  im)
END WHILE
END
```

B. SPEECH TO TEXTUAL ANALYSIS

```
PROCEDURE: Text Analysis
INPUT: train data, test data.tsv
OUTPUT: result, output
VARIABLES: header=0, quoting=3
BEGIN
Read #training/testing
Download text data sets
Initialize an empty list [ ]
WHILE i<= train[review]
   DO
     Append from kaggleword2vecutility to empty
    Create a bag of words from training test
     Initialize count vectorizer object
           i) Fit the model
           ii) Transfer the training data
     Display training random forest
C repeat empty list [ ]
DISPLAY parsing test data
END WHILE
WHILE i <=  test[review]
   DO
     Append from kaggleword2vecutility to empty
     Get bag of words for the test
     DISPLAY predicting test labels
    Copy results to pandas data frame
    DISPLAY  work reuse to bag_of_words.CVS
END WHILE
END
```

VI. RESULTS AND DISCUSSION

Emotion prediction through face expression:

By using the proposed algorithm and the executable code we classified the human emotion by the following process/steps:

a. Input images and pre-processing the given image
b. Expression compound
c. Network structure

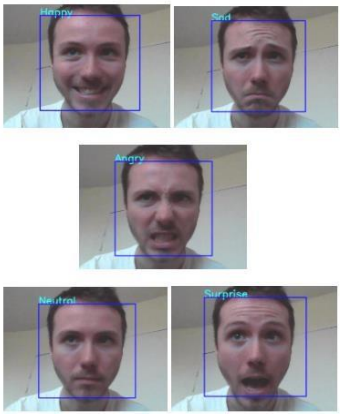
ISBN 978-1-7281-6202-7

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

d. Emotion analysis
e. Evolution of emotion space
f. Emotion recognition from the image

Where the evolution of human emotion is done with the help of the emotion space and the emotion is recognized on live facial recognition.

Images of live-Face expression recognition



Emotion prediction through Voice to Text:

By using this proposed algorithm and the executable code we classified the human speech by the following steps :

a. Speech input
b. Feature extraction
c. Feature selection
d. Classifier
e. Reorganized output

Here, the speech analysis of human is nothing but an application of pattern recognition where the pattern is derived as such energy, pitch, MFCC is mapped using the classifiers namely ANN, HMM, SVM and lots more.

By using our Executable code we can classify the output of our input speech to text and classify the text output to the human emotion as the following:

1. POSITIVE EMOTION (empty, neutral, love, fun, enthusiasm)

2. NEGATIVE EMOTION (sad, worry)

The Emotion in Text data set by kaggleword2vecutility https://www.kaggle.com/c/word2vec-nlp-tutorial/data .By using the above data set 250 data has to be implemented in pycharm or python platform to get a result.

REFERENCE

[1] S. Morishima, H. Harashima. "Emotion space for analysis and compound offacial expression", Proceedings of 1993 2nd IEEE International Workshop onRobot and Human Communication, 1993.

[2]Tengfei Song,Wenming Zheng,Cheng Lu,Yuan Zong,Xilei Zhang,Zhen Cui ," MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition", Proceedings on IEEE Access , Vol-7, Page(s):12177-12191(2019).

[3]Seyedehsamaneh Shojaeilangari ,Wei-Yun Yau , Karthik Nandakumar ,Jun Li ,Eam Khwang Teoh," Robust Representation and Recognition of Facial Emotions Using Extreme Sparse Learning", Proceedings on IEEE Transactions on Image Processing, Vol-24, Issue:7, Page(s):2140-2152(2015).

[4]Xin Kang , Fuji Ren , Yunong Wu," Exploring latent semantic information for textual emotion recognition in blog articles", Proceedings on IEEE/CAA Journal of Automatica Sinica, Vol-5, Issue:a6, Page(s):2011-2016 (2018).

[5]M. S. Likitha ,Sri Raksha R. Gupta ,K. Hasitha ,A. Upendra Raju, "Speech based human emotion recognition using MFCC", Proceedings on 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET),Page(s):2257-2260 (2017).

[6] Pojala Chiranjeevi ,Viswanath Gopalakrishnan ,Pratibha Moogi, "Neutral Face Classification Using Personalized Appearance Models for Fast and Robust Emotion Detection", Vol-24, Issue:9, Page(s)2701-2711 (2015).

[7] Chloé Clavel ,Zoraida Callejas," Sentiment Analysis: From Opinion Mining to Human-Agent Interaction", Proceedings on IEEE Transactions on Affective Computing, Vol-7, Issue:1, Page(s):74-93 (2015).

[8] Soujanya Poria ,Navonil Majumder ,Rada Mihalcea ,Eduard Hovy, "Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances", Proceedings on IEEE Access, Vol-7, Page(s):100943-100953 (2019).

[9] Panagiotis C. Petrantonakis ,Leontios J. Hadjileontiadis, " Emotion Recognition From EEG Using Higher Order Crossings", Proceedings on IEEE Transactions on Information Technology in Biomedicine, Vol-14, Issue:2, Page(s):186-197 (2010).

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

[10] Ruhul Amin Khalil,Edward Jones,Mohammad Inayatullah Babar,Tariqullah Jan,Mohammad Haseeb Zafar,Thamer Alhussain, “Speech Emotion Recognition Using Deep Learning Techniques: A Review”, Proceedings on IEEE Access, Vol- 7, Page(s): 117327 – 117345 (2019).

[11] Kah Phooi Seng,Li-Minn Ang, “Video Analytics for Customer Emotion and Satisfaction at Contact Centers”, Proceedings on IEEE Transactions on Human-Machine Systems, Vol-48, Issue:3, Page(s):266-278(2017).

[12] Maria Teresa Valderas Yamuza , Juan Bolea , Michele Orini , Pablo Laguna , Carlos Orrite , Montserrat Vallverdu , Raquel Bailon, “Human emotion characterization by heart rate variability analysis guided by respiration”, Proceedings on IEEE Journal of Biomedical and Health Informatics, Page(s):1 (2019).

[13] R. Aishwarya, Yogitha. R, Kiruthiga. V, “Smart Road Surface Monitoring with privacy preserved scheme for
vehicle crowd Sensing”, Proceedings of Journal of Computational and Theoretical Nanoscience, Volume 16, Number 8, August 2019, pp. 3204-3210(7).

[14] M. S. Asish and R. Aishwarya, "Cyber Security at a Glance," 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2019, pp. 240-245.

[15] Muppala Sunny Chowdhary, R. Aishwarya, Basam pravallika, “Improvement of E-Governance and M-Governance in polyglot Countries with Digital Etymology exploitation Indo-Aryan linguistics” Proceedings on International Journal of Scientific Research in Computer Science Applications and Management Studies, ISSN 2319 – 1953, Volume 8, Issue 2 (March 2019).

[16] The website: http:\\liwc.wpengine.com

[17] P. Musa, F. A. Rafi and M. Lamsani, "A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition," *2018 Third International Conference on Informatics and Computing (ICIC)*, Palembang, Indonesia, 2018, pp. 1-6.

[18] T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang and Z. Cui, "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," in IEEE Access, vol. 7, pp. 12177-12191, 2019.

[19] Q. Mao, Q. Zhu, Q. Rao, H. Jia and S. Luo, "Learning Hierarchical Emotion Context for Continuous Dimensional Emotion Recognition From Video Sequences," in IEEE Access, vol. 7, pp. 62894-62903, 2019.

BIBILOGRAPHY



Ms R. Aishwarya, She is working as an Assistant Professor in the stream of Computer Science and Engineering at Sathyabama Institute of science and technology. She is an expert in Fog Computing and Cloud Computing. Her research interest is Cloud computing, Internet of Things. She is a Gold medallist in her PG in the year of 2016-2018.



Master Muppala Sunny Chowdhary graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering have excellence in this stream. His research interest is Etymology, Android Applications in real life, Neural Networking, Data Analysis.



Master Akula Abinay graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering have excellence in this stream. His research interest is Cyber Security, Android Applications in real life, Neural Networking, Data Analysis and Ethical Hacking.



Master Kethamukkala Phanindhra graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering have excellence in this stream. His research interest is Cyber Security, Android Applications in real life, Neural Networking, Data Analysis.



Ms.Nadupuru Yamini graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering have excellence in this stream. Her research interest is Cyber Security, Android Applications in real life, Neural Networking, Data Analysis.



Ms.Reema rose toppo graduating from the beloved university Sathyabama Institute of Science and Technology in the stream of Computer Science and Engineering have excellence in this stream. Her research interest is Cyber Security, Android Applications in real life, Neural Networking, Data Analysis, Andriod development.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:00:06 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Cloud computing based Drug Delivery System in the Manufacturing of Drugs and Online Purchase

Dr.P.Asha[1]
[1]Asst.Prof, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai.
ashapandian225@gmail.com

Polu Hemath Kiran Reddy[2]
[2]Student, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai.
polukiranreddy@gmail.com

P.Venkata Subba Manikanta[3],
[3]Student, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai.
manidarling9007@gmail.com

Kuna Sri Sai[4]
[4]Student, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai.
kunasrisai@gmail.com

Abstract—**This work aims to provide accommodation for the rising demand for manufacturing and delivery system of a pharmaceutical company, in the data-driven model of smart factories and industries in the health care sector. The data-driven model of smart factories has three-layered architecture. The deployment, perception and execution layers in the smart industry sector are needed for a delivery system. This paper provides an efficient and flexible source for drug-packaging demand-based production. If certain medicines are rarely available in nearby places, then it could be purchased online. But some medicines should be stored in specific temperatures or some which should be stored in a vacuum environment. So, for the medicine to arrive from foreign countries and reach us without losing their efficiency, the procedures and precautions of their packing of medicines are analyzed to make the drug function effectively without losing its originality. Also need to know their Manufacturer details, dates and the way of ingestion specified in the label, which may differ sometimes. It should be checked for childproof too. All these are carried out using the proposed work to analyze and control different mechanisms according to packaging and label standards.**

Keywords—Pharmaceutical production; Smart factories; Drug-Packaging flow; Healthcare Industry.

I. INTRODUCTION

The Pharmaceutical Industries are facing scourge in the production and manufacturing of products in this demand-based market trend. Thus they are in a challenging situation that needs a more dynamic control system. The traditional design which is made for limited product production was not able to satisfy new product types and cannot adapt to small-batch varieties for manufacturing in rigid patterns. The control system involved in manufacturing needs to be flexible enough to meet the standards of the market in the industry.

To improve reconfigurable production, the significance of the control system is to be enhanced. The functional aspects and production features are essential for achieving greater potential.

Developed technologies like automated machines and smart applications on cloud computing and cyber-security systems are useful for pharmaceutical production purposes. This industry is concentrating on providing safe and qualified products rich in quality for market needs. Thus a control strategy for such manufacturing and production is essential for such purposes and this new era of the modern decade is made possible.

II. REVIEW ON EXISTING WORK

Wang et al. propose the techniques used in smart factory creation like 3D-printed technology and physical system software. These techniques are used for improving production by certain procedures to meet the market demands in the smart industry [1-3].

An approach for management is proposed, to bring up management to develop sustainable lean land. It consists of a system of measures and methods capable of such potential to achieve that goal [3-4]. The author proposes a structured architecture for reconfigurable manufacturing of pharmaceutical products. This study gives complete ideas for smart factory applications via reconfiguration management and its perspective is credible [5-8].

Another work summarizes the modern methods used in the pharmaceutical industry for raising the charts of production. Modern technology including cloud computing and security methods are improved and can be used for this purpose [9-12].

The Existing system uses the MySQL database as a backend for storing data. This is a major limitation of this system. Whenever data load increases this system cannot handle and breaks the operation. The system can be improved by replacing that backend with the Hadoop tool [13-16].

A. Drawbacks

- The processing time is high.
- Data loss cannot be undone
- Data security cannot be assured
- Flexibility and speed can be improved

III. PROPOSED METHODOLOGY

The proposed system uses the Hadoop tool for processing data and is far better than the existing system. The database provided by the Hadoop tool does not have any limitation over the database and can add any other number of machines to the

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:43:22 UTC from IEEE Xplore. Restrictions apply.

cluster. The technique used by this tool makes this system more useful for its purpose of usage. It uses Hadoop for processing the data.

A. *Advantages of the Proposed System*

- High throughout
- Less maintenance cost
- Fast response time
- Joins, Partitions and bucketing Techniques are used
- No data loss problem.
- Efficient data processing.

Map-Reduce has,

1. Shuffle stage.
2. Reduce stage.
3. Map stage.

Map stage: Its job was to process the input data. Normally, the input data was in the form of a file or index and it is stored in the Hadoop. The file is proceeding to the mapper function line after line. The mappers process the data and create several small blocks of data.

Reduce: This platform is used for the fusion of the stages and the reduction stage. The Reducerjob is to process the data that comes to the map. After the process of the data, it makes a new set of output, which will be stored in the Hadoop file system.

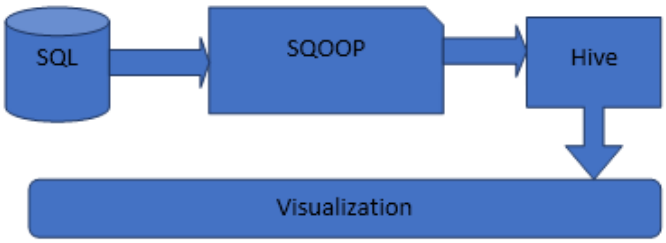


Fig. 1. Data Flow diagram

The process flow of the proposed system starts from HDFS and ends at Hive. Hive is a data warehouse used for storing and analyze data for the process. The data will be gathered for analysis by using the SQOOP Import command. The data analyzing process is performed by HIVE. The flow continues by extracting meaningful data and removing the unclenched data because the HIVE can only process structures data and effectively analyze it (Fig 1).



Fig. 2. Storage

The existing system used a MySQL database. Thus by importing that backed up data from the server and can send it to processing. The data collection from that server is done by the SCOOP import command to draw out that data to HDFS. The data stored there is sent for processing to the HIVE and gets analyzed there (Fig 2).

The proposed system has 3 main portions,

1. Converts the .csv files send to the MySQL
2. Transfer the MySQL to the Hadoop using Sqoop
3. Analysis of data using hive
4. Pre-processing
5. Visualization of the data

B. *Convert .csv file and send to MYSQL*

Here the medical data is collected from the manufacturing industries or even from the online, which are in the form of .csv file. A CSV file is imported. Examine the data and clean them. This CSV file is then carried to MySQL. All those backup data which have stored in MYSQL is imported by the use of Sqoop commands to Hadoop Distributed File System. Now all these data are stored in HDFS where it is ready to be handled by the hive (Fig 3).

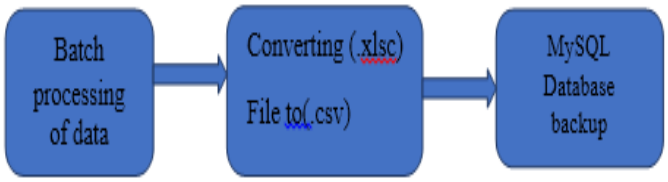


Fig. 3. Data-processing

C. *Transfer MYSQL to Hadoop using Sqoop*

The SQL data is modified and transformed into the Hadoop. Sqoop is an application that is used to import the SQL data to the Hadoop. Hadoop is a software that is used for storing and processing enormous data. It contains two things one is HDFS and map-reduce. Sqoop is a command-line interface application for transfer of Drug data between relational databases (MySQL) and Hadoop (Fig.4).



Fig. 4. SQOOP

In Above Fig 4, MySQL database contains an aggregative dataset that has to import it to Hadoop Files System using Sqoop. Drug data can be moved into a Hadoop file system or Hive from MySQL and then it will bring about the java classes. In going before cases, the continuous run of data was from Relational Data Base Management System to HDFS. Using the "export" tool can import data from the Hadoop files system to the Relational Database Management System. Before operating export, Sqoop gets table metadata from the MySQL database. Thus, a table has to be created firstly, with the necessary metadata.

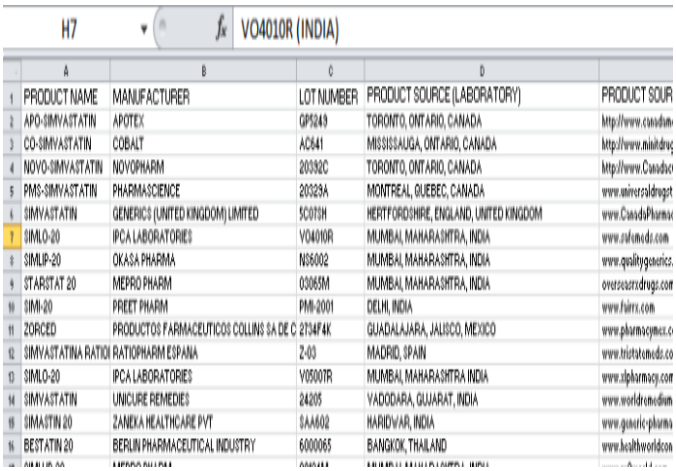
Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:43:22 UTC from IEEE Xplore. Restrictions apply.

D. Analysis of data with Hive

Hive is used to find a particular data analysis.it provides an instrumental process to project the architecture on to the data and query data is used SQL language and HiveQL. Here in HIVE, only structured data can be processed to analyze. By extracting only, the meaningful data and neglecting unclenched data, the data can be analyzed more effectively by use of Hive.

IV. RESULTS AND DISCUSSION

First, the data is imported as a .csv file. MySQL is logged in and import the data then creating the database by product name, manufacturing date, maintenance, cost, etc. and put the data into the Hadoop using Sqoop it converts into the hive for the analysis. If you wanted to do a partition of data and can show particular data. This data is downloaded and send them to the R system for the graph representation.



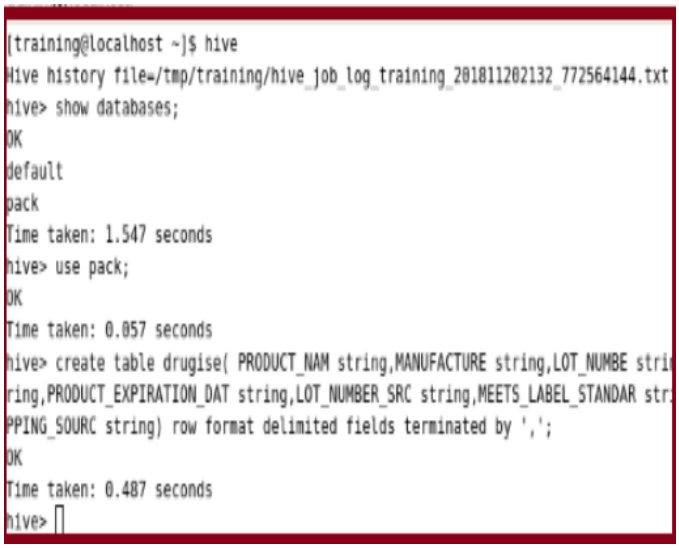
H7 VO4010R (INDIA)

	A	B	C	D	
1	PRODUCT NAME	MANUFACTURER	LOT NUMBER	PRODUCT SOURCE (LABORATORY)	PRODUCT SOUR
2	APO-SIMVASTATIN	APOTEX	GP5249	TORONTO, ONTARIO, CANADA	http://www.canadian
3	CO-SIMVASTATIN	COBALT	AC641	MISSISSAUGA, ONTARIO, CANADA	http://www.mainlydrug
4	NOVO-SIMVASTATIN	NOVOPHARM	20332C	TORONTO, ONTARIO, CANADA	http://www.Canadian
5	PMS-SIMVASTATIN	PHARMASCIENCE	20329A	MONTREAL, QUEBEC, CANADA	www.universaldrugst
6	SIMVASTATIN	GENERICS (UNITED KINGDOM) LIMITED	5C07SH	HERTFORDSHIRE, ENGLAND, UNITED KINGDOM	www.CanadaPharma
7	SIMLO-20	IPCA LABORATORIES	VO4010R	MUMBAI, MAHARASHTRA, INDIA	www.safemeds.com
8	SIMLIP-20	OKASA PHARMA	N05002	MUMBAI, MAHARASHTRA, INDIA	www.qualitygenerics.
9	STARSTAT 20	MEPRO PHARM	03065M	MUMBAI, MAHARASHTRA, INDIA	overseasrxdrugs.com
10	SIMI-20	PREET PHARM	PMI-2001	DELHI, INDIA	www.fairrx.com
11	ZORCED	PRODUCTOS FARMACEUTICOS COLLINS SA DE C	2734F4K	GUADALAJARA, JALISCO, MEXICO	www.pharmacymex.c
12	SIMVASTATINA RATIO	RATIOPHARM ESPANA	Z-03	MADRID, SPAIN	www.tristatemeds.co
13	SIMLO-20	IPCA LABORATORIES	V05007R	MUMBAI, MAHARASHTRA INDIA	www.xlpharmacy.com
14	SIMVASTATIN	UNICURE REMEDIES	24205	VADODARA, GUJARAT, INDIA	www.worldremedium
15	SIMASTIN 20	ZANEKA HEALTHCARE PVT	SAA602	HARIDVAR, INDIA	www.generic-pharma
16	BESTATIN 20	BERLIN PHARMACEUTICAL INDUSTRY	6000065	BANGKOK, THAILAND	www.healthworldcon

Fig. 5. Medical data

Collect the medicine data and store in the XML format where it transforms into the .xlsc to a CSV file it consists of thousands of data of a data set that is known as mock data. Connects to the MySQL server and login into the MySQL using MySQL commands creating the database and after that, the tables are created to store the data. The CSV data is imported into the MySQL table (Fig 5).

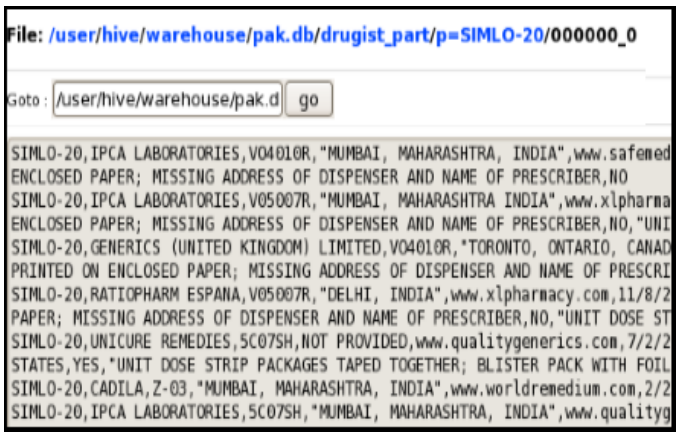
Hive uses a charge-based optimizer. The input of data is being analyzed using the hive tool based upon the drug name or even the packing or the manufacturing dates. It used to determine and elaborate on the problems a data that has been present in it.it uses the map-reduce algorithm and optimal query. Analyzing is a tool for queries that shows your data required (Fig 6).



```
[training@localhost ~]$ hive
Hive history file=/tmp/training/hive_job_log_training_201811202132_772564144.txt
hive> show databases;
OK
default
pack
Time taken: 1.547 seconds
hive> use pack;
OK
Time taken: 0.057 seconds
hive> create table drugise( PRODUCT_NAM string,MANUFACTURE string,LOT_NUMBE stri
ring,PRODUCT_EXPIRATION_DAT string,LOT_NUMBER_SRC string,MEETS_LABEL_STANDAR str
PPING_SOURC string) row format delimited fields terminated by ',';
OK
Time taken: 0.487 seconds
hive>
```

Fig. 6. Hive analysis

A hive is a tool. Where in which it divides of tables into the related part is known as the partition of hive.it is either done by columns like date product name, manufacturing date, etc. it has more partition keys to identify a particular partition.it organizes the data in the hive.in Fig 7. That it displays the partitioned data.



File: /user/hive/warehouse/pak.db/drugist_part/p=SIMLO-20/000000_0

Goto : /user/hive/warehouse/pak.d go

```
SIMLO-20,IPCA LABORATORIES,VO4010R,"MUMBAI, MAHARASHTRA, INDIA",www.safemed
ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO
SIMLO-20,IPCA LABORATORIES,V05007R,"MUMBAI, MAHARASHTRA INDIA",www.xlpharma
ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO,"UNI
SIMLO-20,GENERICS (UNITED KINGDOM) LIMITED,VO4010R,"TORONTO, ONTARIO, CANAD
PRINTED ON ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRI
SIMLO-20,RATIOPHARM ESPANA,V05007R,"DELHI, INDIA",www.xlpharmacy.com,11/8/2
PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO,"UNIT DOSE ST
SIMLO-20,UNICURE REMEDIES,5C07SH,NOT PROVIDED,www.qualitygenerics.com,7/2/2
STATES,YES,"UNIT DOSE STRIP PACKAGES TAPED TOGETHER; BLISTER PACK WITH FOIL
SIMLO-20,CADILA,Z-03,"MUMBAI, MAHARASHTRA, INDIA",www.worldremedium.com,2/2
SIMLO-20,IPCA LABORATORIES,5C07SH,"MUMBAI, MAHARASHTRA, INDIA",www.qualityg
```

Fig. 7. Drug dataset in a hive

In the hive, the table data sets are more manageable and the concepts of hive partitioning provide a way of isolation of the hive table data into multiple files/directories. It uses a hash function in the bucketing process and column will always be stored in the same bucket (Fig 8). It displays the bucketing of datasets that are utilized in this process. A hash function maps any arbitrary data size into a fixed-sized one. The hash values and their hash tables help for effective bucketing.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:43:22 UTC from IEEE Xplore. Restrictions apply.

```
SIMLO-20,IPCA LABORATORIES,VO4010R,"MUMBAI, MAHARASHTRA, INDIA",www.safemeds.com,6/
ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO
SIMLO-20,IPCA LABORATORIES,VO5007R,"MUMBAI, MAHARASHTRA INDIA",www.xlpharmacy.com,1
ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO,"UNIT DOSE S
SIMLO-20,GENERICS (UNITED KINGDOM) LIMITED,VO4010R,"TORONTO, ONTARIO, CANADA",www.s
PRINTED ON ENCLOSED PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO
SIMLO-20,RATIOPHARM ESPANA,VO5007R,"DELHI, INDIA",www.xlpharmacy.com,11/8/2016,12/4,
PAPER; MISSING ADDRESS OF DISPENSER AND NAME OF PRESCRIBER,NO,"UNIT DOSE STRIP PACK
SIMLO-20,UNICURE REMEDIES,5C07SH,NOT PROVIDED,www.qualitygenerics.com,7/2/2017,3/7/
STATES,YES,"UNIT DOSE STRIP PACKAGES TAPED TOGETHER; BLISTER PACK WITH FOIL BACKING
SIMLO-20,CADILA,Z-03,"MUMBAI, MAHARASHTRA, INDIA",www.worldremedium.com,2/23/2016,1,
SIMLO-20,IPCA LABORATORIES,5C07SH,"MUMBAI, MAHARASHTRA, INDIA",www.qualitygenerics.
UNITED STATES,YES
```

Fig. 8. Bucketing dataset

The detection of errors caused by noise or other damage done inside the factory and manufacturing of medicine during transmission from the transmitter to the receiver.in the manufacturing of medicine goes under the inspection at the source. If things are done correctly 99 percent of the time, the 1% will be the error that can be reduced using these methods and analysis of the medicine and with exports and imports also. This method is used to prevent these errors, like labelling and missing products during the manufacturing of drugs (Fig 9).

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.6439  -1.6125  -1.3947   0.1101   9.8470

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.552397   0.106475  52.148  < 2e-16 ***
Websites     0.004448   0.001600   2.780  0.00544 **
PName        0.004079   0.001593   2.562  0.01042 *
Lab         -0.023874   0.002173 -10.986  < 2e-16 ***
Label_std    0.075021   0.051826   1.448  0.14775
Package_std  0.008886   0.072607   0.122  0.90259
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 11.70691)

    Null deviance: 563429  on 47999  degrees of freedom
Residual deviance: 561862  on 47994  degrees of freedom
A: 254315
```

Fig. 9. Error detection of standards and error values

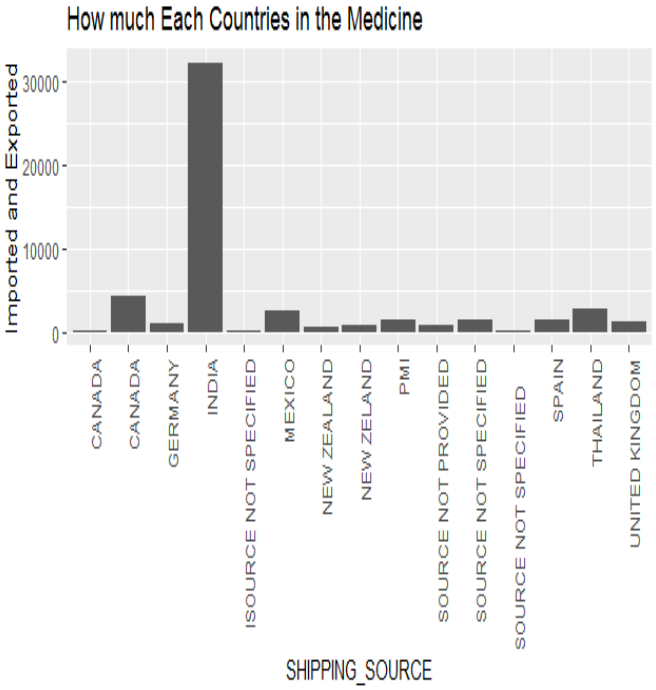


Fig. 10. Shipping of medicine

Fig 10 displays a large number of data volumes in the chart form or the bar graph. It displays the image of the product that is labeled and not labeled or even the import and exports of medicines by which the medicine supplied to other countries.

V. CONCLUSION

The work proposes on drug manufacture, drug production and drug delivery with a maintenance system. Online purchasing makes this more useful for everyone who has a smart application. Smart factories are implemented to manufacture and produce drugs. Therefore, the proposed study on drug attempts at helping patient, pass therapy and improvise compliance. It is a reconfigurable approach to produce, deliver and maintenance of an effective drug delivery system which is a boon to the pharmaceutical industry. This helps the pharmaceutical industry to improve its service with a quick response for various medical needs in the market.

REFERENCES

[1] S. Wang, J. Wan, D. Li, and C. Zhang, "Implementing smart factory of Industrie 4.0: An outlook," International Journal of Distributed Sensor Networks, vol. 12, no. 1, 2016, Art. ID. 3159805.

[2] J. Wan, S. Tang, Z. Shu, D. Li, and S. Wang, "Software-Defined Industrial Internet of Things in the Context of Industry 4.0," IEEE Sensors Journal, vol. 16, no. 4, pp. 7373-7380, 2016.

[3] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, "Towards the smart factory for Industrie 4.0: A self-organized multi-agent system assisted with big data-based feedback and coordination," Computer Networks, vol. 101, pp. 158–168, 2016.

[4] V. C. Gungor, et al, "Smart grid technologies: Communication technologies and standards," IEEE transactions on industrial informatics, vol. 7, no. 4, pp. 529-539, 2011.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:43:22 UTC from IEEE Xplore. Restrictions apply.

[5] M. Armbrust, A. Fox, R. Griffith, et al. "A view of cloud computing," Communications of the ACM, vol. 53, no. 4, pp. 50-58, 2010.

[6] J. Wan, D. Zhang, S. Zhao, L.T. Yang, and J. Lloret, "Context-aware Vehicular Cyber-Physical Systems with Cloud Support: Architecture, Challenges, and Solutions," IEEE Communications Magazine, vol. 52, no. 8, pp. 106–113, 2014.

[7] Sabina Jeschke, Christian Brecher, Housing Song, and Danda Rawat, "Industrial Internet of Things: Cyber manufacturing Systems," Cham, Switzerland: Springer, pp. 1-715, 2017.

[8] Housing Song Danda Rawat, Sabina Jeschke, and Christian Brecher, "Cyber-Physical Systems: Foundations, Principles and Applications," Boston, MA: Academic Press, pp. 1-514, 2016.

[9] P. Renna, "Capacity reconfiguration management in reconfigurable manufacturing systems," International Journal of Advanced Manufacturing Technology, vol. 46, no. 1-4, pp. 395-404, 2010.

[10] Ankayarkanni, B., J. Albert Mayan, and J. Aruna. "Support Vector Machine for Effective Robust Visual Tracking." Journal of Computational and Theoretical Nanoscience 16, no. 8 (2019): 3571-3575.

[11] K Haripriya, G Harshini, L. Sujihelen, "Survey on Efficient Automated Toll System for License Plate Recognition Using Open CV", 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR), pp. 1-6, 2018.

[12] Pandian, Asha, Bharathi Varadharajulu, Albert Mayan John, and Prem Jacob. "A Comprehensive View of Scheduling Algorithms for Mapreduce Framework in Hadoop", Journal of Computational and Theoretical Nanoscience 16, no. 8 (2019): 3582-3586.

[13] P.Asha, Dr.T.Jebarajan, "SOTARM: Size of Transaction based Association Rule Mining Algorithm", Turkish Journal of Electrical Engineering and Computer Sciences, 25(1), pp. 278-291, 2017. DOI: 10.3906/elk-1406-75.

[14] Repaka, A.N., Ravikanti, S.D., Ramya G. Franklin(2019)," Design and implementing heart disease prediction using naives Bayesian",Proceedings of the International Conference on Trends in Electronics and Informatics, ICOEI 2019,PP-292-297.

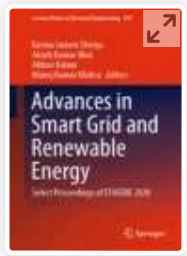
[15] K. Ashokkumar, Dr.C.Chandrasekar, MKS.Akshya, 2015, Efficient Method For Secure Key Matching Process Of Large Data Set Integration In Grid Computing, Research Journal Of Applied Sciences, Engineering And Technology, Vol.10(3), pp.288-292, ISSN. 2040-7467.

[16] Asha P., Lahari T., Kavya B, "Comprehensive Behaviour of Malware Detection Using the Machine Learning Classifier", In: Soft Computing Systems. ICSCS 2018. Communications in Computer and Information Science, vol 837. Springer, Singapore, 2018.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:43:22 UTC from IEEE Xplore. Restrictions apply.

Springer Link

Search Log in



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 651–657

Choosing an Ephemeris for Low Earth Orbit (LEO) Satellite Application

T. Omkara Chari, T. Manikanta, G. Nagarajan, T. Prem Jacob & A. Pravin

Conference paper | First Online: 05 January 2021

319 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

An epic versatile call affirmation control algorithm for sight and sound low orbit satellite systems was proposed. In light of continuous call dropping likelihood of goal cell, the algorithm can save transfer speed for handoff calls consolidating likelihood limit technique and reasonable utility distribution plot. Recreation results show that the current algorithm presents acceptable new call blocking likelihood, extraordinarily decreases handoff call dropping

likelihood, even though ensures the huge transfer speed use. A fruitful satellite strategic activities can be accomplished by the labor and the control system in amicability. The proposed system create high accuracy look into grade positions for planets and Earth satellites. It utilizes highly exact planetary ephemeris documents that are created by NASA JPL and are utilized for deciding the exact area of nearby planetary group objects. The Development Ephemeris (DE) documents De200, De405, De418, De421, and De430 are furnished with free flyer and contain information for Sun, Moon, the eight significant planets, and Pluto.

Keywords

IOT **Green data collection**

Low Earth Orbit (LEO) satellite **Equitable utility**

Quality of service

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_66
- Chapter length: 7 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Price includes VAT (India)

- ISBN: 978-981-15-7511-2
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7714-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7510-5
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. T. Reid, A. Neish, T. Walter, P.K. Enge, Leveraging commercial broadband LEO constellations for navigation, in: *Proceedings of page 11 of 11 54 the ION GNSS+2016* (Institute of Navigation, Portland, Oregon, USA, 2016), pp. 2300–2314

2. M. Joerger, J. Neale, B. Pervan, Iridium/Gps Carrier phase positioning and fault detection over wide areas, in *Proceedings of the ION GNSS 2009* (Institute Of Navigation, Savannah, Georgia, USA, 2009), pp 1371–1385

3. M. Centenaro, L. Vangelista, A. Zanella, M. Zorzi, Longrange communications in unlicensed bands: the rising stars in the IoT and smart city scenarios. IEEE Wirel. Commun. **23**(5), 60–67 (2016)

4. R. Christina, G.L. Shajan, B. Ankayarkanni, Cart—A statistical model for predicting QOE using machine learning in smartphones, in *IOP Conference Series: Materials Science and Engineering*, vol. 590, no. 1 (Iop Publishing, 2019), p. 012001

5. H.S. Dhillon, H. Huang, H. Viswanathan, Wide-area wireless communication challenges for the

Internet of Things. IEEE Commun. Mag. **55**(2), 168–174 (2017)

6. R. Barbatei, A. Skavhaug, T.A. Johansen, Acquisition andrelaying of data from a floating wireless sensor node using anunmanned aerial vehicle, in *International Conference On Unmannedaircraft Systems* (ICUAS) (2015), pp. 677–686

7. A. Zolich, J.A. Alfredsen, T.A. Johansen, K.R. Skøien, A Communication bridge between underwater sensors and unmanned vehicles using a surface wireless sensor network-design and validation, in *Oceans 2016-Shanghai* (2016), pp. 1–9

8. M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, R. Prasad, Satellite communications supporting internet of remote things. IEEE Internet of Things J. **3**(1), 113–123 (2016)

9. https://www.Inmarsat.Com/Service/Isatdatapro/

10. S.C. Mana, M. Saipriya, S.K. Sangeetha, Identification of land document duplication and black money transaction using big data analytics, in *2019 Fifth International Conference on Science Technology Engineering and*

Mathematics (ICONSTEM), Chennai, India (2019), pp. 114–118

11. K. Sangeetha, P. Vishnuraja, D. Deepa, Stable clustered topology and secured routing using mobile agents in mobile ad hoc networks. Asian J. Inf. Technol. **15**(23), 4806–4811 (2016)

12. K.K. Thyagharajan, R.I. Minu, Prevalent color extraction and indexing. Int. J. Eng. Technol. **5**(6), 4841–4849 (2013)

13. M.J. Neely, E. Modiano, C.E. Rohrs, Dynamic powerallocation and routing for time-varying wireless networks. IEEE J. Sel. Areas Commun. **23**(1), 89–103 (2005).

14. N. Accettura, M.R. Palattella, G. Boggia, L.A. Grieco, A. Dohler, Decentralized traffic aware scheduling for multi-hoplow power lossy networks in the Internet of Things, in *IEEE 14th International Symposium and Workshops on a World of Wireless, Mobile and Multimedia Networks (WOWMOM)* (2013), pp. 1–6

15. S. Tian, W. Dai, R. Liu, J. Chang, G. Li, System using hybrid LEO-GPS satellites for rapid resolution of integer cycle ambiguities. IEEE

Trans. Aerosp. Electron. Syst. **50**(3), 1774–1785 (2014)

16. X. Luo, G. Liu, H. Huang, Buffer capacity-constrained epidemic routing model in mobile ad-hoc networks, in *IEEE 17th International Conference On Computational Science and Engineering (CSE)*, 1443–1448 (2014)

17. P. Gonizzi, G. Ferrari, V. Gay, J. Leguay, Data dissemination scheme for distributed storage for IoT observation systems at large scale. Inf. Fusion **22**, 16–25 (2015)

18. C.-T. Cheng, N. Ganganath, K.-Y. Fok, Concurrent data collection trees for IoT applications. IEEE Trans. Ind. Inf. **13**(2), 793–799 (2017).

19. D.U. Nandini, S. Divya, A literature survey on various watermarking techniques, in *2017 International Conference On Inventive Systems and Control* (ICISC) (IEEE, 2017), pp. 1–4

20. S.L. Shabu, C. Jayakumar, Detection of brain tumor by image fusion using genetic algorithm. Res. J. Pharmaceut. Biol. Chem. Sci. **7**(5), 510–516 (2016)

21. H. Wu, J. Li, H. Lu, P. Hong, A two-layer caching model forcontent delivery services in satellite-terrestrial networks, in *IEEE Global Communications Conference (Globecom)* (2016), pp. 1–6

22. M. Divya, M.D. Kamalesh, Recovery of watermarked image from geometrics attacks using effective histogram shape based index. Ind. J. Sci. Technol. **9**(44), 1–6 (2016)

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

T. Omkara Chari, T. Manikanta, G. Nagarajan, T. Prem Jacob & A. Pravin

Corresponding author

Correspondence to T. Omkara Chari .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Omkara Chari, T., Manikanta, T., Nagarajan, G., Prem Jacob, T., Pravin, A. (2021). Choosing an Ephemeris for Low Earth Orbit (LEO) Satellite Application. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_66

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7511-2_66

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Change Detection of Images Based on Multivariate Alteration Detection Method

[1]Jeberson Retna Raj, [2]Senduru Srinivasulu
[1,2]*Department of IT, School of Computing*
[1.2]Sathyabama Institute of Science and Technology, Chennai, India
[1]jebersonretnarajr@gmail.com

Abstract— **Advancement of Satellite technology gives rapid response to the complex problems and decision making. Due to the impact of climate change, loss of rainfall, drought, risen of sea level are some of the consequences. Recently, the cyclone Fani hits oddisa severely and the livelihood of the people is in standstill. The satellite images of before and after Fani shows the severity of the cyclone. This paper, we introduced three methods which includes Large Scale Mean Shift (LSMS), Multivariate Alteration Detection (MAD) and Local statistic Feature extraction methods are used to detect the changes in the heterogeneous images. The satellite images of before and after the cyclone Fani is considered for processing and the severity is estimated. The satellite images are calibrated and preprocessed for removing the speckle noises. With the application of regression prediction algorithm used to tune the accuracy of segmentation and change detection process. The model is tested with the satellite image and the results are very promising. This approach can also be applied in medical images to detect the changes in features, which helps the physician to diagnose the decease.**

Keywords— Multivariate Alteration Detection, Satellite images, Change Detection

I. INTRODUCTION

Climate change is one of the key issues among countries for it negative influences make impact on human lives and environment. The impact that includes rising sea level, drought, loss of rain fall, ozone depletion changes in coastal landscape etc. Due to industrialization, the emission of CO2 level is increased the record high of 441 rpm, which is never before for the past 80000 years. Various reasons cited as deforestation, urbanization, air pollution etc. The human factors involved in polluting all the three dimensions of environment that includes land, air and water. The major eco system that includes aquatic, terrestrial, desert and mountain eco systems are affected by human. One of the disastrous is unseasoned cyclone. Cyclone Fani is one such distrastrous hits odissa coast between bahramphur and puri on may 8th 2019 with a speed of more than 150 kms. The daily life of people was devastated and millions of people stayed in the relief camps. The satellite images released by NASA shows that the kind of worst disastrous before and after Fani cyclone. In this paper, the satellite picture of Fani before and after is taken and the changes are detected and demarcated. The multivariate alteration Detection and local statistic Feature Extraction algorithm used to extract the affected region.

II. Literature Survey

Dalmia et al proposed a feature extraction method for detecting changes in the satellite images. The Structural Phase Congruency Histogram method is proposed to build the structural properties of corner and edges features of the image. The Linear Discriminant Analysis method is used to reduce the diamentionality of the image and the changes between images are estimated with the support of ANN algorithm[1]. Akansha et al proposed a novel method to detect the changes in satellite images. The spectral changes detected by change vector analysis and the statistical changes identified by finding similarity index based on kullback leibler distance. The wavlet fusion techniques is used to combine the spectral change and similarity index in the images[2]. Wei Cui et al proposed a change detection algorithm based on extracting the multi direction feature vector, multi scale log file is extracted by Non-subsampling contourlet transform (NSCT) and with the help of k-means algorithms the changed region is detected[3]. Barazzetti presents a change detection technique based on object to object comparison between two satellite images. The affine transform is used to detect the changes and the silver is removed with the support of geometrical techniques. In essence, the silver is introduced in the satellite images due to variable off-nadir angles, weak geo-location and variable off-nadir angles. This limitation is overcome in this method[4]. Xi et al presented a spatiotemporal cube model for analyzing and change detection of satellite images. The spatio temporal cube model consist of the characteristics which includes spatio temporal correlation, scalability and heterogeneity[5]. They introduced a multi resolution segmentation method for segment the homogeneous spatio temporal cubes[6]. Jeppesen et al presented a cloud detection algorithm for detecting cloud in a complex satellite image. They introduced a deep learning algorithm named after RS-Net used to detect the cloud coverage in the image[7].

Ning Lv et al prposed a feature learning method for change detection in satellite image using stacked Contractive Auto Encoder method(sCAE). This method works based on finding the temporal difference in pixel level and linear iterative clustering algorithm used to generate super pixels as the change sample for learning input. The sCAE network make use of the change sample as input and generate the change map as the output. The approach is compared with the Principal Component Analysis and markov random fields[8]. Yangyang Li et al proposed a deep learning model for detecting the changes in heterogeneous satellite images. The CNN algorithm is used to capture the spatial characteristics of raw data from the image and generate false label by using fuzzy clustering algorithm. With the selected samples as input, the CNN is trained and the final classification result is obtained [9]. Maoguo et al proposed a deep learning approach for change detection in satellite image. The approach make use of joint fuzzy classier for labeled feature of the two input image and with the help of RBM network used for pretraining and the fine tuning method used to detect the changes in the images[10].

978-1-7281-5197-7/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:03:49 UTC from IEEE Xplore. Restrictions apply.

III. Methodology

Change detection in heterogeneous satellite image is paramount important for analyzing and interpreting the changes in Region Of Interest (ROI) of earth surface. The satellite image of oddisa before Fani and after Fani is considered for this work. The illumination of light effect at night before and after the cyclone is released by NASA open source images. The image is co-registered for loading the image for reading and calibrated for setting the world coordinates for further processing. Generally, the satellite image by default contains speckle noise due to the position of sensor of satellite for capturing the land surface and the atmospheric effect. The speckle noise is to be removed by the preprocessing algorithm. The noise free image is used as the input for segmentation and change detection. The proposed system architecture is shown in figure 1.

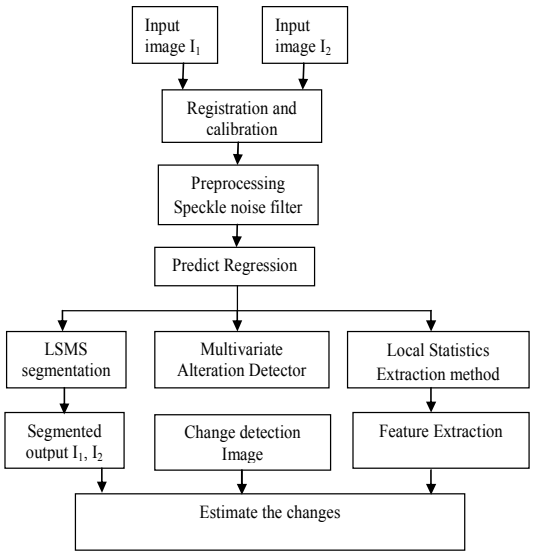


Figure 1: Architecture of the proposed Change detection method

The two heterogeneous images of I1 and I2 is considered as a input for processing. The images are co-registered and calibrated for fit in the tool for process the image. The speckle noise is removed from the image and the resultant image is given as an input for predict regression. To improve the accuracy of segmentation and multivariate alteration detection algorithm, the output of the predict regression is used. The Large Scale Meanshift Segmentation (LSMS) is used to segment the ROI image detail with other part of the image. The multivariate alteration detection algorithm used to detect the changes in the two satellite images. Finally the severity of changes can be estimated with the prediction algorithm.

The work consist of the following steps:

Step 1: identify the satellite images for change detection.

Step 2: satellite images are registered and calibrated for processing.

Step 3: preprocessing algorithm for removal of speckle noise.

Step 4: predict regression.

Step 5: Applying with LSMS segmentation method, Multivariate Alteration Detection algorithm and local statistical method

Step 6: Estimate the percentage of changes and quantify it.

A. Identify the satellite images

The satellite images of before and after Fani cyclone is considered for implementation. NASA released the open source LANSAT images of land surface of oddisa during night light before and after Fani is publically available at their website. These images consist of six bands and each band can be processed extensively.

B. Image registration and Calibration

The LANDSAT images have by default having different dimensions. For processing these images using any of the images processing a tool box which requires the fixation of coordinate positions and geometric correction made with the defined Geo-coordinates. This step ensures for applying any image processing algorithm for extraction of information from the image.

C. Prediction of regression

The prediction of the image is making use of the statistic information of image file such as mean and standard deviation and the output image is predicted. The input model file of image detail is given as a input and the image mask is created based on the threshold greater than 0 is used for prediction process.

D. Multivariate Alteration Detector (MAD) algorithm

This algorithm is used to detect the changes in multispectral satellite images. The two heterogeneous images with same coordinate position are taken as a input and the relation between change map of different channel is computed. The six different features of change map like variance, mean are calculated for different bands and the difference of image is constructed.

E. Large Scale Mean Shift Segmentation (LSMS) method

The method produces the labeled image whose neighborhood pixel radius is below the given threshold is grouped under the same cluster and the image is segmented. The spatial radius value is set as 5 and range radius value is set as 15 for clustering the pixels.

F. Local Statistics Extraction method

The local statistics extraction method used to extract the features of image based on the input given. The mean, variance and standard deviation of neighborhood pixels are considered and the similar pixel with the class labels identified and classified.

G. Estimate the changes

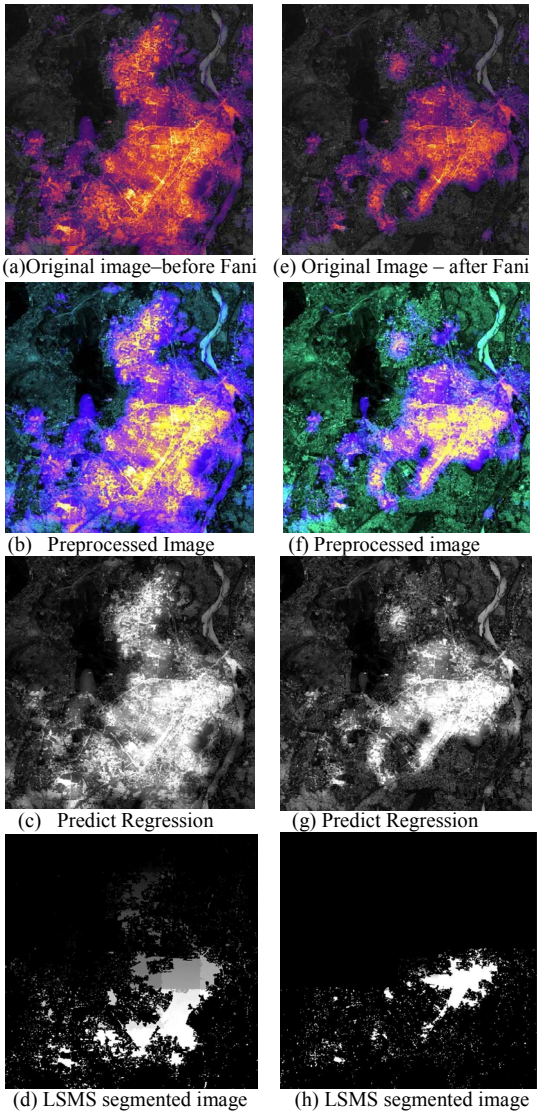
The random forest algorithm is used to forecast the percentage of changes between the images and with the help of this prediction, we can quantify the severity of disastrous.

IV. Experimental Result

The LANSAT images of before and after Fani cyclone devastated between bahrampur and poori the state government of Oddisa, India is considered as an input for our work. The images I1 and I2 of before and after cyclone is

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:03:49 UTC from IEEE Xplore. Restrictions apply.

registered, calibrated and geometrically corrected for processing. The image contains the speckle noise due to the atmospheric effect and sensor angles which capture the land surface. The filtered image is further used to predict the regression and the new single band image is obtained for segmentation. The multivariate alteration method is used to detect the changes in the two images. Figure 2(a) shows the original satellite image of oddisa during the light emission of land surface at night. Figure 2(b) shows the preprocessed image after removing the speckle noise. Figure 2(c) shows the predicted image of the application with predict regression method. Figure 2(d) shows the LSMS segmented image.



(a)Original image–before Fani (e) Original Image – after Fani

(b) Preprocessed Image (f) Preprocessed image

(c) Predict Regression (g) Predict Regression

(d) LSMS segmented image (h) LSMS segmented image

Figure 2 (a) – (d) before Fani and (e) – (h) after Fani

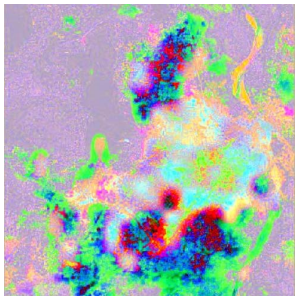
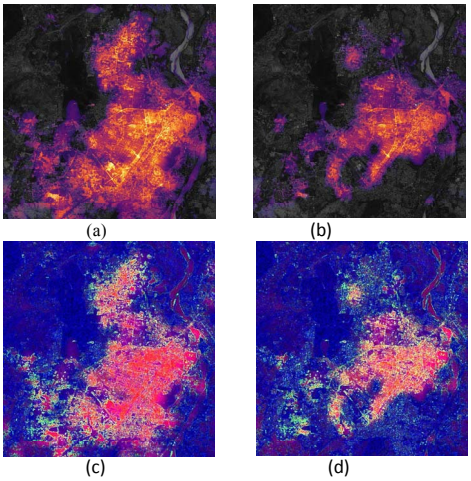


Fig 2 (i) Multivariate alteration detection

Figure 2(e) shows the original LANDSAT image of earth surface aftermath of the cyclone Fani hits oddisa. The electricity connections affected areas are shown in the photograph at night. Figure 2(f) shows the preprocessed image and figure 2(g) shows the predicted image. The LSMS segmented image of after Fani is shown in figure 2(h). Figure 2(i) shows the output of multivariate alteration detection method.



(a) (b)

(c) (d)

Figure 3 (a) and (b) original image of before and after Fani
(c) and (d) output of Local statistic extraction method

Table 1 shows that the comparison of Mean Square Error and Root Mean Square error between two images compared and listed.

TABLE 1 : COMPARISON OF MSE AND PSNR VALUES

Method	MAE	PSNR
LSMS segmentation	11.63	14.51
Local statistics extraction method	19.27	15.08

Table 1 shows the comparison of MAE, MSE and PSNR values between the outputs of the two images. The lower these values are the higher the performance. Hence, LSMS segmentation is one of the good techniques for extracting the affected region in the image

V. CONCLUSION

The change detection of satellite images of Fani is processed and successfully interpreted by the three algorithms which include Large Scale Mean Shift segmentation, Multivariate Change Detection algorithm and Local statistical extraction algorithms. The images are loaded and preprocessed with the help of speckle filter and the noise free image is generated. The regression prediction algorithm used to filling the missing values and the output of this step used to increase the accuracy. The performance of these algorithms with the given architecture is compared and the results are analyzed. These algorithms can be applied with the similar kind of applications in change detection domain. In future, the severity of damages, loss of crop, vegetation, agriculture can be estimated with the help of other deep learning algorithms. The same algorithm can be applied in medical images to detect the changes in cells and tissues which help the user to take decisions.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:03:49 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Dalmiya, C. P., Santhi, N., & Sathyabama, B. (2018). A novel feature descriptor for automatic change detection in remote sensing images. The Egyptian Journal of Remote Sensing and Space Science.

[2] Akansha Singh , Krishna Kant Singh, "Unsupervised change detection in remote sensing images using fusion of spectral and statistical indices", The Egyptian Journal of Remote Sensing and Space Sciences, PP: 21 (2018) 345–351

[3] Cui, W., Jia, Z., Qin, X., Yang, J., & Hu, Y. (2011). Multi-temporal Satellite Images Change Detection Algorithm Based on NSCT. Procedia Engineering, 24, 252–256.

[4] Barazzetti, L. (2016). Sliver Removal in Object-Based Change Detection from VHR Satellite Images. Photogrammetric Engineering & Remote Sensing, 82(2), 161–168.doi:10.14358/pers.82.2.161

[5] Xi, W., Du, S., Wang, Y.-C., & Zhang, X. (2019). A spatiotemporal cube model for analyzing satellite image time series: Application to land-cover mapping and change detection. Remote Sensing of Environment, 231, 111212. doi:10.1016/j.rse.2019.111212

[6] Jeppesen, J. H., Jacobsen, R. H., Inceoglu, F., & Toftegaard, T. S. (2019). A cloud detection algorithm for satellite imagery based on deep learning. Remote Sensing of Environment, 229, 247–259. doi:10.1016/j.rse.2019.03.039

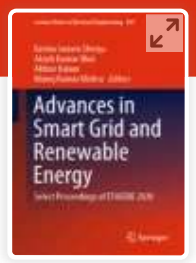
[7] Li, J., Wang, C., Wang, S., Zhang, H., Fu, Q., & Wang, Y. (2017). Gaofen-3 sea ice detection based on deep learning. 2017 Progress in Electromagnetics Research Symposium - Fall (PIERS - FALL).doi:10.1109/piers-fall.2017.8293267

[8] Ning Lv, Chen Chen, Tie Qiu, Arun Kumar Sangaiah, Deep Learning and Superpixel FeatureExtraction Based on Contractive Auto encoder for Change Detection in SAR Images, IEEE Transactions on Industrial Informatics, VOL. 14, NO. 12, DECEMBER 2018, PP:553-5538

[9] Yangyang Li, Cheng Peng, Yanqiao Chen, Licheng Jiao, Linhao Zhou, Ronghua Shang, "A Deep Learning Method for Change Detection in Synthetic Aperture Radar Images", IEEE Transactions on Geoscience And Remote Sensing, Vol. 57, No. 8, August 2019

[10] Maoguo Gong, Jiaojiao Zhao, Jia Liu, Qiguang Miao, Licheng Jiao, "Change Detection in Synthetic Aperture Radar Images Based on Deep Neural Networks", IEEE Transactions On Neural Networks And Learning Systems, VOL. 27, NO. 1, January 2016, PP:125-138

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:03:49 UTC from IEEE Xplore. Restrictions apply.



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 529–537

Blockchain Integration in House Voting System

Hemanth Chebrolu, Saikumar Buragadda, A. Viji Amutha Mary ✉ & Jancy

Conference paper | First Online: 05 January 2021

305 Accesses | **1** Citations

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Developing a secured electronic democratic framework assures the safety and decency of existing democratic, without compromising in straightforwardness, and flexibility. This has only been a trial for current frameworks. The electronic democratic has risen being a substitution for paper–pen-based democratic to decrease the duplications and irregularities. In this existing current framework, ballot-based voting is available; yet at the same time, there is no framework to stay away from proxy casting and recasting is actualized.

We do not have a choice to see our made choice also. There is no security right now. In the proposed framework, a novel blockchain-dependent e-casting ballot framework is provided. This electronic democratic framework is improved from the existing framework by using blockchain structures. The portions of impediments are used to assess the enhancement made from the existing system. In altering the undertaking part, we coordinate aadhar card connected versatile number for OTP age, at exactly that point the voter can make the choice; this framework forestalls throwing and re-throwing of intermediaries.

Keywords

I-voting **E-voting** **Electronic democratic** **Blockchain-based voting**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_52
- Chapter length: 9 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33
Price includes VAT (India)

- ISBN: 978-981-15-7511-2
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7714-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 279.99

Price excludes VAT (India)

- ISBN: 978-981-15-7510-5
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. X. Li, et al., in *What are Intelligent Contracts? A Guide for Intelligent Contracts for Beginners* (Blockgeeks, 2016). At:

https://blockgeeks.com/guides/intelligentcontracts/

2. T.P. Jacob, T. Ravi, Optimal regression test case prioritization using genetic algorithm. Life Sci. J. **10**(3), 1021–1033 (2013)

3. H. Pan et al., A multi-part ballot based E-voting system with clash attack protection (2015)

4. A. Barnes, B. Christopher, P. Thomas, Digital voting using Blockchain technology (2016). https://www.Economist.com/sites/defaults/files/plymouth.pdf

5. O.M. Olaniyi, et al., Secure vote today (2016). Available at: https://www.lawfareblog.com/secure-vote-today

6. T. Kunioka, G.M. Woller, In (a) democracy we trust: social and economic determinants of support for democratic procedures in Central and Eastern Europe'. J. Soc. Econ. **28**(5), 577–596 (1999)

7. T. van der Meer, 'In what we trust? A multi-level study into trust in parliament as an evaluation of state characteristics.' Int. Rev. Administ. Sci. **76**(3), 517–536 (2010)

8. G. Nagarajan, R.I. Minu, B. Muthukumar, V. Vedanarayanan, S.D. Sundarsingh, Hybrid genetic algorithm for medical image feature extraction and selection. Proc. Comput. Sci. **85**, 455–462 (2016)

9. J. Andrews, Evaluation of various compiler optimization techniques related to mibench benchmark applications (2013)

10. D. Basin, H. Gersbach, A. Mamageishvili, L. Schmid, O. Tejada, Election security and economics: it's all about eve, in *Proceedings of the International Joint Conference on Electronic Voting* (2017), pp. 1–28

11. P. Bevelander, R. Pendakur, Electoral participation as a measure of social inclusion for natives, immigrants, and descendants in Sweden. Technical report (2008), p. 33

12. S. Wolchok et al., Security analysis of India's electronic voting machines, in *Proceedings of the 17th ACM Conference on Computer and Communications Security* (2010), pp. 1–14

13. A. Viji Amutha Mary, S. Justin Samuel, Automated trinity based web data extraction for simultaneous comparison. Contemp. Eng. Sci. **8**(11), 491–497 (2015) (HIKARI Ltd., Scopus)

14. J. Stern, Vote—voting for a mobile world. Accessed: 31 July 2018 [Online]. Available: https://votem.com/

15. M. Pilkington, 11 Blockchain technology: principles and applications, in *Research Handbook on Digital Transformations* (2016), p. 225

16. G. Gabison, Policy considerations for the blockchain technology public and private applications. SMU Sci. Tech. Rev. **19**, 327 (2016)

17. M. Feldhofer, S. Dominikus, J. Wolkerstorfer, Strong authentication for RFID systems using the AES algorithm, in *Cryptographic Hardware and Embedded Systems* (2004), pp. 357–370

18. Y. Stein, H. Primo, Programmable data encryption engine for advanced encryption standard algorithm. U.S. Patent 7 508 937 B2 (2009). Accessed: 1 Aug 2018 [Online]. Available: https://patents.google.com/patent/US7508937B2/en

19. B. Shehzad, K.M. Awan, M.I.-U. Lali, W. Aslam, Identification of patterns in the failure of software projects. J. Inf. Sci. Eng. **33**(6), 1465–1480 (2017)

20. A.M. Abdullatif, B. Shahzad, A. Hussain, Evolution of social media in scientific research: a case of technology and healthcare professionals in Saudi Universities. J. Med. Imag. Health Inform. **7**(6), 1461–1468 (2017)

21. B. Shahzad, Identification of risk factors in large scale software projects: a quantitative study. Int. J. Knowl. Soc. Res. **5**(1), 1–11 (2014)

22. A.B. Shahzad, A. Said, Application of quantitative research methods in identifying software project factors. Int. J. Inf. Technol. Elect. Eng. **1**(1), 30–33 (2012)

23. A. Viji Amutha Mary, M. Paul Selvan, Christy, Public auditing for secure cloud storage using MD5 algorithm. Int. J. Recent Technol. Eng. (IJRTE) **8**(3) (2019). ISSN: 2277-3878

24. M.A. Chowdary, M. Kundan, D.A.V.A. Mary, Effective credit card forgery prevention using multilevel authentication. IOP Conf. Ser. Mater. Sci. Eng. **590**(1), 012021 (2019) (IOP Publishing)

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Hemanth Chebrolu, Saikumar Buragadda, A. Viji Amutha Mary & Jancy

Corresponding author

Correspondence to A. Viji Amutha Mary .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Chebrolu, H., Buragadda, S., Viji Amutha Mary, A., Jancy (2021). Blockchain Integration in House Voting System. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_52

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-7511-2_52

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)
SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Springer Link

Search Log in



Proceedings of the 2nd International Conference on Communication, Devices and Computing pp 399–404

Biometric-Based Three-Tier Microservice Architecture for Mitigating the Fraudulent Behaviour

S. Prayla Shyry

Conference paper | First Online: 17 December 2019

536 Accesses | **4** Citations

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 602)

Abstract

Though monolithic systems have tremendous advantages, it suffers from its own drawbacks like independency and debugging time. So the paradigm has shifted to build the applications with microservice architecture. The microservice is a part of the service-oriented architecture that decomposes the monolithic architecture into different services. All the divided services perform heterogeneous activity, and they are interfaced with proper application programming interface (API). Microservice is highly resilient; hence, when one service fails, it does not have an impact on other services. Traditional systems use monolithic

architecture which fails in their design pattern and dependency. In this paper, to avoid duplicity in passport, a novel biometric-based monolithic architecture is developed for authentication system. All the physiological and behavioural biometrics of a person are stored in the repository and compared during the authentication phase. All these services are created as microservice and interfaced with proper APIs which improves parallel processing.

Keywords

Application programming interface **Biometric** **Microservice**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-0829-5_39
- Chapter length: 6 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 117.69
Price includes VAT (India)

- ISBN: 978-981-15-0829-5
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only

- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 139.99

Price excludes VAT (India)

- ISBN: 978-981-15-0831-8
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 199.99

Price excludes VAT (India)

- ISBN: 978-981-15-0828-8
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. Villamizar, M., Garcés, O., Castro, H., Verano, M., Salamanca, L., Casallas, R., Gil, S.: Evaluating the monolithic and the microservice architecture pattern to deploy web applications in the cloud. 583–590. IEEE (2015). 978-1-4673-9464-2/15/

2. Gan, Y., Delimitrou, C.: The architectural implications of cloud micro-services. 1–4 (2018)

3. Villamizar, M., Garces, O., Ochoa, L., Castro, H., Salamanca, L., Verano, M., Casallas, R., Gil, S., Valencia, C., Zambrano, A., Lang, M.: Infrastructure cost comparison of running web applications in the cloud using AWS lambda and monolithic and microservice architectures. In: 2016 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, pp. 179–182

4. Nginx: Adopting microservices at netflix: lessons for architectural design (2015)

5. Buyya, R.: Cloud computing: the next revolution in information technology. In: 2010 1st International Conference on in Parallel Distributed and Grid Computing (PDGC), pp. 2–3, Oct 2010

6. Lewis, J., Fowler, M.: Microservices (2014)

7. Decomposing twitter: adventures in service-oriented architecture. In: 2014 Microservices Workshop: Why, What, and How to Get There. http://www.slideshare.net/adriancockcroft/microservices-workshop-craft-conference

8. Courjon, D., Sarayeddine, K., Spajer, M.: Scanning tunneling optical microscopy. Opt. Commun. **71**(1-2), 23–28 (1989)

Author information

Authors and Affiliations

School of Computing, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

S. Prayla Shyry

Corresponding author

Correspondence to S. Prayla Shyry .

Editor information

Editors and Affiliations

Department of Electronics and Communication Engineering, National Institute of Technology Durgapur, Durgapur, West Bengal, India

Dr. Sumit Kundu

Department of Electronics and Communication Engineering, National Institute of Technology Karnataka, Mangalore, Karnataka, India

Dr. U. Shripathi Acharya

Department of Electronics and Communication Engineering, Haldia Institute of Technology, Haldia, West Bengal, India

Dr. Chanchal Kr. De

Department of Electronics and Communication Engineering, Haldia Institute of Technology, Haldia, West Bengal, India

Surajit Mukherjee

Rights and permissions

Reprints and Permissions

Copyright information

© 2020 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Prayla Shyry, S. (2020). Biometric-Based Three-Tier Microservice Architecture for Mitigating the Fraudulent Behaviour. In: Kundu, S., Acharya, U., De, C., Mukherjee, S. (eds) Proceedings of the 2nd International Conference on Communication, Devices and Computing. Lecture Notes in Electrical Engineering, vol 602. Springer, Singapore. https://doi.org/10.1007/978-981-15-0829-5_39

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-15-0829-5_39

Published	Publisher Name	Print ISBN
17 December 2019	Springer, Singapore	978-981-15-0828-8

Online ISBN	eBook Packages
978-981-15-0829-5	Physics and Astronomy Physics and Astronomy (R0)

Not logged in - 14.139.184.28
Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

SCIENTIFIC REPORTS
nature research

Check for updates

OPEN

Bioconversion of chitin and concomitant production of chitinase and N-acetylglucosamine by novel *Achromobacter xylosoxidans* isolated from shrimp waste disposal area

Kumaran Subramanian[1,4,11], Balamurugan Sadaiappan[2,11], Wilson Aruni[1,9], Alagappan Kumarappan[3], Rajasekar Thirunavukarasu[4], Guru Prasad Srinivasan[5], Selvaraj Bharathi[6], Prasannabalaji Nainangu[6], Pugazhvendan Sampath Renuga[7], Anandajothi Elamaran[5,8], Deivasigamani Balaraman[5] & Mahendran Subramanian[10]✉

Marine pollution is a significant issue in recent decades, with the increase in industries and their waste harming the environment and ecosystems. Notably, the rise in shellfish industries contributes to tons of shellfish waste composed of up to 58% chitin. Chitin, the second most ample polymer next to cellulose, is insoluble and resistant to degradation. It requires chemical-based treatment or enzymatic hydrolysis to cleave the chitin polymers. The chemical-based treatment can lead to environmental pollution, so to solve this problem, enzymatic hydrolysis is the best option. Moreover, the resulting biopolymer by-products can be used to boost the fish immune system and also as drug delivery agents. Many marine microbial strains have chitinase producing ability. Nevertheless, we still lack an economical and highly stable chitinase enzyme for use in the industrial sector. So we isolate a novel marine bacterial strain *Achromobacter xylosoxidans* from the shrimp waste disposal site using chitin minimal medium. Placket–Burman and central composite design statistical models for culture condition optimisation predicted a 464.2 U/ml of chitinase production. The culture conditions were optimised for maximum chitinase production recording up to 467 U/ml. This chitinase from the *A. xylosoxidans* was 100% active at an optimum temperature of 45 °C (withstand up to 55 °C) and pH 8 with 80% stability. The HPLC analysis of chitinase degraded shellfish waste reveals a major amino acid profile composition—arginine, lysine, aspartic acid, alanine, threonine and low levels of isoleucine and

[1]School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600119, India. [2]Plankton Ecology Laboratory, CSIR- National Institute of Oceanography, Panaji, Goa 403004, India. [3]Sharjah Institute for Medical Research, University of Sharjah, Sharjah, UAE. [4]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600119, India. [5]Centre for Advanced Studies in Marine Biology, Annamalai University, Chidambaram, Tamil Nadu 608502, India. [6]Department of Microbiology, Sri Sankara Arts and Science College, Enathur, Tamil Nadu 631561, India. [7]Department of Zoology - DDE, Annamalai University, Chidambaram, Tamil Nadu 608002, India. [8]Central Aquaculture Genetics Laboratory, Rajiv Gandhi Centre for Aquaculture, Karaimedu, Tamil Nadu 609109, India. [9]School of Medicine, Loma Linda University, Loma Linda, CA 92350, USA. [10]Department of Bioengineering, Department of Computing, Imperial College London, London, SW72AZ, UK. [11]These authors contributed equally: Balamurugan Sadaiappan and Kumaran Subramanian. ✉email: m.subramanian@imperial.ac.uk

methionine. These chitinase degraded products and by-products can be used as supplements in the aquaculture industry.

Chitin is made up of N-acetylglucos-amine through a β-(1 → 4) glycosidic bond to assemble poly(β-(1 → 4)-N-acetyl-d-glucosamine (GlcNaC) and it is by far the second most ample biopolymer present in nature. Which is insoluble and resistant to degradation, and it requires chitinase (EC 3.2.1.14) or closely related enzymes for its breakdown. About 10^{12}–10^{14} tons of chitin is produced as waste per year[1] and the major problem for the aquaculture industries is to dispose of this waste. If it is not adequately disposed of or degraded, this waste may cause potential peril to the natural environment and biodiversity[2]. Out of the shellfish industry waste, about 20–58% contains chitin as dry weight[3].

The best reaction to this issue is to use these losses as inexhaustible crude material for the production of valuable items that can be used in aquaculture, textile, and medical industries[4]. Chitin is in many cases firmly bound with different mixes, for example, lipids, protein and calcium carbonate[5]. Bioconversion of the squandered chitinous waste to valuable chitooligosaccharides involves procedures such as deproteinisation, demineralisation or hydrolysis. These procedures have previously been performed with solid corrosives and bases that incorporates low yields, high levels of expense and consumption issues[6]. The potential substitute for solving this problem is chitin waste management by chitinase[7].

Glycosyl-hydrolase proteins (EC 3.2.2.14), otherwise known as chitinases can cleave the 1, 4 bonds of the N-acetylglucosamine units, hydrolyse the chitin to chitooligosaccharides [(GlcNAc)n] and act as a catalyst in chitin debasement[8]. Chitosan oligomers (a group of N-deacetylated chitin with different degrees of deacetylation) and other derivatives from chitin degradation by either compound or enzymatic hydrolysis are used in the several fields of the pharmaceutics, horticulture, biotechnology and waste administration. Chitinase is naturally produced by microbes, fungi, plants, insects and animals[9]. Among them, chitinase produced by microbes has received increasing consideration, and potentially fill two needs: (1) they decrease the ecological perils of waste administration, and (2) can increase the worth of the degraded products[10]. Consequently, marine bacterial chitinases are now considered as one of the potential enzymes for this application.

Many marine bacterial genera were reported to produce chitinase like *Alcaligenes*, *Bacillus*, *Vibrio*, *Acinetobacter* and *Streptomyces rimosus*[11]. However, most naturally occurring marine bacteria cannot produce high quantities of chitinase with maximum activity and stability. It is essential to identify, isolate a potential bacterial strain and optimise the culture conditions to maximise the production in a cost-effective manner. As Baas Becking and Beijerincken[12] mentioned "everything is everywhere, but, the environment selects" most bacterial genera thriving in the shrimp waste disposal site would have the ability to degrade chitin. So, the shrimp waste disposal site should be a potential source for identifying and isolating a suitable chitinase producing bacteria.

For culture condition optimisation, Response Surface Methodology (RSM) is the most broadly used technique[13]. Plackett–Burman (PB) and central composite design (CCD) are multi-purpose examining tools comprising statistical methods for producing empirical models that assist in understanding enzyme kinetics and for optimising the manufacturing conditions of expensive products such as enzymes. The PB design and central composite design were the widely used statistical experimental design in many biotechnological applications[14], enzyme manufacturing[15], biomass production[16] and ethanol production[17] because of its simplicity and most common saturated design[18]. The unbiased estimation of main effects with the smallest variance can be done using saturated fractional factorial designs. In addition to this, the model is orthogonal, so the impact of individual variables will not interfere with the other variable interactions[19]. Herein, we isolated bacterial strains that produce chitinase from a shrimp waste disposal site. Next, we optimised the culture components using RSM to increase the production of chitinase. Later, the amino acid profile of the chitinase degraded products was determined and discussed.

Materials and methods

Colloidal chitin preparation. About 5 g of chitin was mixed with 30 ml of HCl acid (35.5%) and incubated overnight at 4 °C. The colloidal chitin was precipitated by slowly adding 250 ml of chilled ethanol (50%), with constant stirring at 4 °C and left for overnight. The colloidal chitin was centrifuged at 10,000 g for 20 min, and sterile distilled water was used to wash the pellets until the pH value is neutral.

Isolation of chitinase producing bacteria. The soil samples were collected from five different places of the shrimp and crab waste disposable area near Port Novo landing centre situated on the south-east coast of India (Parangipettai—Lat. 11° 29′; long. 79° 46′ E). The 1 g of soil sample was serially diluted and plated on chitin minimal agar plates (Colloidal chitin, 1%; K_2HPO_4, 0.3 g; NaCl, 4 g; KH_2PO_4, 0.3 g; $MgSO_4{\cdot}7H_2O$, 0.5 g; Agar, 1.5% per litre). The plates were incubated at 30 °C for 3 days. The bacterial colony showing the maximum zone of chitin clearance was selected for further studies.

Culture condition for the production of chitinase. A 50 ml of chitin minimal medium (CMM) containing 1% (w/v) colloidal chitin, 0.7% (w/v); KH_2PO_4, 0.3% (w/v); K_2HPO_4, 4% (w/v); NaCl, 0.5% (w/v); $MgSO_4{\cdot}7H_2O$ and 0.5% (w/v) peptone and pH 7.2 was dispensed in a 250 ml flask and inoculated with 1 ml of young seed culture. The flask was incubated at 37 °C for 3–5 days. Further to incubation, the supernatant was separated via centrifugation of culture broth at 8000 g for 10 min (Sigma laboratory centrifuge 4K15, Chennai, India), and was used as the crude enzyme.

Chitinase assay. 2.5 ml of 1% colloidal chitin (substrate) was mixed with 2.5 ml of PBS, and then 0.5 ml crude enzyme was added; next, the flask was incubated at 45 °C for 1 h. The reaction mixture was added with dinitrosalicylic (DNS) acid and was placed within a boiling water bath for 10 min to stop the reaction. The supernatant was collected by centrifuge, and the reduced sugar was estimated at 540 nm (UV spectrophotometer). Using the GlcNAc standard curve, under assay condition, the quantity of enzyme that yields 1 µmol of reducing sugar/ minute can be defined as one unit of chitinase.

Characterisation of isolates. The biochemical and morphological characteristics of chitin degrading isolates were characterised, and the strains were identified using Bergey's manual of determinative bacteriology[20]. The potential bacterial strain was also characterised by 16S rDNA sequencing.

Molecular characterisation. Genomic DNA was extracted from the bacterial isolates by the phenol–chloroform method. The isolated DNA was visualised using 0.8% agarose gel and quantified using NanoDrop 1000 (Thermo Fisher Scientific, Wilmington, DE, USA). The 16S rDNA sequence of the bacterial strain was amplified by PCR using universal bacterial primers 27F and 1492R and GeNei™ PCR Master Mix (Genei, Bangalore, India). A thermal cycler (Genei, Bangalore, India) was used for the PCR reaction – 94 °C for 5 min (initial denaturation), next step involves 94 °C for 1 min (denaturation), 53 °C for 30 s (annealing), 72 °C for 90 s (elongation)—35 cycles and 72 °C for 7 min (final extension)[21].

Phylogenetic analysis. The Genei PCR purification kit (Genei, Bangalore, India) was used in the purification of the amplified PCR products. The sequence of 16s rDNA amplified product was obtained by an automated sanger sequencer (Bioserve, Hyderabad, India). The sequence was edited using BioEdit ver. 7.9[22] followed by a BLAST search in the NCBI public database to identify the closest sequences. The Phylogenetic tree was drawn in the Mega 6.0 version using the neighbour-joining method[23]. The partial 16S rDNA sequence of bacterial strain was deposited in the GenBank database, and accession numbers were received upon submission to the NCBI database.

Extraction and purification of chitinase. The crude enzyme was obtained by centrifuging the production medium at 8000 g for 10 min, followed by precipitation of enzymes at different concentration of [50–80% (w/v)] ammonium sulfate slowly added into the supernatant under constant stirring for 30 min at 4 °C. Next, the protein precipitate was collected by centrifugation at 10,000 g for 30 min at 4 °C and dissolved in Tris–HCl buffer with pH.7.5. Next, the enzymes were dialysed at 4–6 °C against 20 mM Tris–HCl buffer with pH 7.5 and incubated for 24 h, and the buffer was changed at every 8 h interval.

The dialysed enzyme solution was then purified with Sephadex 75 column (20 cm × 1.5 cm) equilibrated with 20 mM Tris–HCl buffer with the pH 8.0. The flow rate of 1 ml/5 min was used to elute the enzyme. The eluted fractions were estimated for the chitinase activity. The fractions showing chitinase activity were mixed together and concentrated by using a fast flow DEAE-Sepharose column (1.6 cm × 20 cm) pre-equilibrated with 20 mM Tris–HCl of pH 8 (10 mmol/l). The bounded chitinase was eluted at a flow rate of 400 µl/min using different gradients of NaCl buffer (0–0.5 mol/l) at 4 °C. The collected chitinase fractions were used as an enzyme source for chitinase assays and protein quantification[24].

Determination of optimal temperature and pH for enzyme activity. The reaction mixture was incubated at different temperatures and tested for chitinase activity to find the optimum temperature for enzyme activity. To determine the thermal stability of the enzyme, 50 mM Tris–HCl buffer-pH 9.0 along with 0.2 ml of the purified enzyme was incubated for 1 h at a temperature range of 35–95 °C with 5 °C interval. The enzyme activity was evaluated by withdrawing the aliquots for every 15 min and tested using standard chitinase assay. To determine the optimal pH for enzyme activity, 0.2 ml of the purified enzyme along with 50 mM Phosphate buffer was incubated at different pH ranging between 4 and 9 and the residual activity was measured using chitinase assay.

Identifying the significant variables for maximum chitinase production. The medium components for maximising the chitinase production by *A. xylosoxidans* were screened by using Plackett–Burman design (PBD)[18]. The medium components chosen for the present study include colloidal chitin, yeast extract, peptone, KH_2PO_4, $FeSO_4$, $ZnSO_4$ and $MgSO_4$, with each component being represented at two levels, low (−1) to high (+1). Table 1 shows the experimental design. The culture medium was prepared accordingly and incubated at 37 °C for 3–5 days. After incubation, the culture filtrate was centrifuged at 8000 g for 15 min. The actual (measured) and estimated (PBD-based) enzyme values are given in Table 1. The following Eq. (1) calculated the effect of individual parameters on the chitinase production:

$$E(X_i) = 2(M_i^+ - M_i^-)/N \quad (1)$$

where E is the effect of the parameter, X_i and M_i^+ and M_i^- are chitinase activity responses, and the number of trials is N.

Optimisation through RSM. The components showing a positive effect on the chitinase production (Table 1), were further optimised using RSM, i.e., we employed a central composite design (CCD) for maximum production of chitinase enzyme. The significant variables for chitinase production, i.e. peptone, chitin, KH_2PO_4 and $MgSO_4$ and a total of 31 experiments were performed with these variables at five different levels (−2, −1,

Blocks	Yeast extract	Peptone g/l	Colloidal chitin (g/l)	KH_2PO_4 g/l	$FeSO_4$ g/l	$ZnSO_4$ g/l	$MgSO_4$ g/l	Enzyme activity (U/ml)	
								Experimental	Predicted
1	1	0.1	50	0.1	0.05	0.05	0.5	218	208.1667
2	1	1	10	1	0.05	0.05	0.1	163.2	178.9
3	0.1	1	50	0.1	0.1	0.05	0.1	212.8	201.8333
4	1	0.1	50	1	0.05	0.1	0.1	179	156.1
5	1	1	10	1	0.1	0.05	0.5	214	198.3
6	1	1	50	0.1	0.1	0.1	0.1	159.2	170.1667
7	0.1	1	50	1	0.05	0.1	0.5	168.4	197.3667
8	0.1	0.1	50	1	0.1	0.05	0.5	203.4	207.1667
9	0.1	0.1	10	1	0.1	0.1	0.1	139.4	129.5667
10	1	0.1	10	0.1	0.1	0.1	0.5	128.2	149.9667
11	0.1	1	10	0.1	0.05	0.1	0.5	200.8	171.8333
12	0.1	0.1	10	0.1	0.05	0.05	0.1	145.2	162.2333

Table 1. Plackett–Burman based experimental design.

Run order	Peptone g/l	Chitin g/l	KH_2PO_4 g/l	$MgSO_4$ g/l	Enzyme activity U/ml	
					Experimental	Predicted
1	0.325	20	0.325	0.2	403	399.5833
2	0.775	20	0.325	0.2	390	385.5417
3	0.325	40	0.325	0.2	377	378.5417
4	0.775	40	0.325	0.2	405	404.25
5	0.325	20	0.775	0.2	383	383.2083
6	0.775	20	0.775	0.2	388	388.9167
7	0.325	40	0.775	0.2	397	394.9167
8	0.775	40	0.775	0.2	437	440.375
9	0.325	20	0.325	0.4	439	434.2083
10	0.775	20	0.325	0.4	386	388.9167
11	0.325	40	0.325	0.4	398	397.9167
12	0.775	40	0.325	0.4	394	392.375
13	0.325	20	0.775	0.4	352	353.5833
14	0.775	20	0.775	0.4	331	328.0417
15	0.325	40	0.775	0.4	347	350.0417
16	0.775	40	0.775	0.4	360	364.25
17	0.1	30	0.55	0.3	372	373.7083
18	1	30	0.55	0.3	375	373.875
19	0.55	10	0.55	0.3	373	377.7083
20	0.55	50	0.55	0.3	397	392.875
21	0.55	30	0.1	0.3	416	421.0417
22	0.55	30	1	0.3	381	376.5417
23	0.55	30	0.55	0.1	407	409.0417
24	0.55	30	0.55	0.5	369	367.5417
25	0.55	30	0.55	0.3	459	462
26	0.55	30	0.55	0.3	463	462
27	0.55	30	0.55	0.3	466	462
28	0.55	30	0.55	0.3	462	462
29	0.55	30	0.55	0.3	459	462
30	0.55	30	0.55	0.3	467	462
31	0.55	30	0.55	0.3	458	462

Table 2. CCD based experimental design and results.

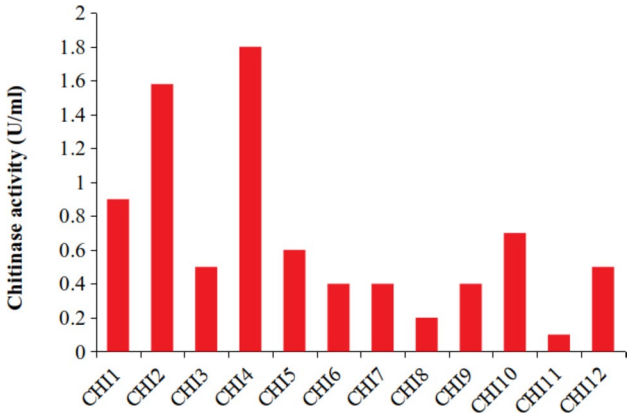


Figure 1. Screening of potential strains of chitinase producing bacteria.

0, +1, and +2). The full experimental plan, including the values of the significant variables, is given in Table 2. The enzyme activity U/ml (response value) was obtained by calculating the average of the triplicate in each trial. Analysis of variance (ANOVA) was performed on the data obtained from RSM on Chitinase production. The response surface regression involves the polynomial equation (2):

$$Y = \beta_0 + \sum_i \beta_{ii} X_i + \sum_i \beta_{ii} X_i^2 + \sum_i \beta_{ij} X_i X_j \quad \tilde{u}'\beta \tag{2}$$

where Y is the prediction, β_0 is the constant term, β_i is the ith linear coefficient, β_{ii} is the ith quadratic coefficient, β_{ij} is the ijth interaction coefficient, and X_i, X_j *are* independent variables. The independent variables were expressed as $X1$, $X2$, $X3$, $X4$, $X5$ and can be represented as the second-order polynomial equation (3):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \beta^{ii} X_i^2 + \sum_i \sum_j^k \beta_j X_i X_{j+\epsilon} \tag{3}$$

Fisher's check was used to evaluate the statistical importance of the equation version and model terms. The quality of fit for the second-order polynomial equation was expressed through the coefficient of determination (R^2) and the adjusted R^2. The fitted polynomial equation (3) was expressed as 3-dimensional surface plots to visualise the relationship between the responses and the utilised variables. Point optimisation approach was deployed to optimise each variable for the maximum reaction[25].

Amino acid analysis. The colloidal chitin was mixed with 500 µl of purified chitinase and incubated at 45 °C for 1 h. Subsequent to incubation, the centrifugation of the reaction mixture was performed at 3500 g for 15 min. 1 N NaOH was used to neutralise the filtered supernatant. The neutralised supernatant was diluted to 100 times of the volume by adding milli-Q water, and 20 µl was injected into HPLC (Shimadzu, USA) containing LC1oAT HPLC pumps. The analysis was carried out with 30% milli-Q water and 70% acetonitrile at a flow rate of 1000 µl/min in 40 °C as mobile phase and detected by SPD-M10A detector with UV absorbance at 250 nm.

Ethical approval. No animal or human studies were carried out by the authors.

Results and discussion

Isolation and characterisation of chitin producing strains. In this present investigation, the isolation and characterisation of potent chitin degrading bacteria from the marine environment were performed. Fourteen bacterial strains were isolated from the shrimp waste disposal site, among the 14 strains one bacterial strain showing maximal chitinase activity (1.8 U/ml, Fig. 1) was selected, and this strain was named as chitinase producing bacterial strain 4 (CHI4). The morphological and biochemical test showed this strain CHI4 was rod-shaped, Gram-negative, motile bacteria, utilising catalase and oxidase and oxidise sugars such as xylose and glucose. The 1091 bp 16S rDNA sequence and blast analysis showed that this bacterial strain was closely similar to *Achromobacter xylosoxidans*. Moreover, 99 per cent similar to the *Achromobacter xylosoxidans* strain Osb (Genbank accession no: MN889379.1) and *Achromobacter xylosoxidans* strain Spa05NA (Genbank accession no: MT052658.1). The accession number in the NCBI database for the deposited 16S rDNA partial sequence was JQ756451. The phylogenetic tree for the selected strains was constructed using the neighbour-joining method (Fig. 2). Earlier, this species was named as *Alcaligenes xylosoxidans*. Later, based on the 16S rDNA sequence, it was renamed to *Achromobacter xylosoxidans*[26]. It is widely found in soil, water, dialysis and chlorhexidine solutions, sometimes present in the gastrointestinal and respiratory tract of humans[27].

In the past, numerous marine bacterial strains were reported to produce chitinase such as *Bacillus cereus* SV1[28], *Alteromonas* sp. O-7[29], *Aeromonas hydrophila*[30], *Pseudoalteromonas* sp. DC14[31], *Pseudoalteromonas* sp.

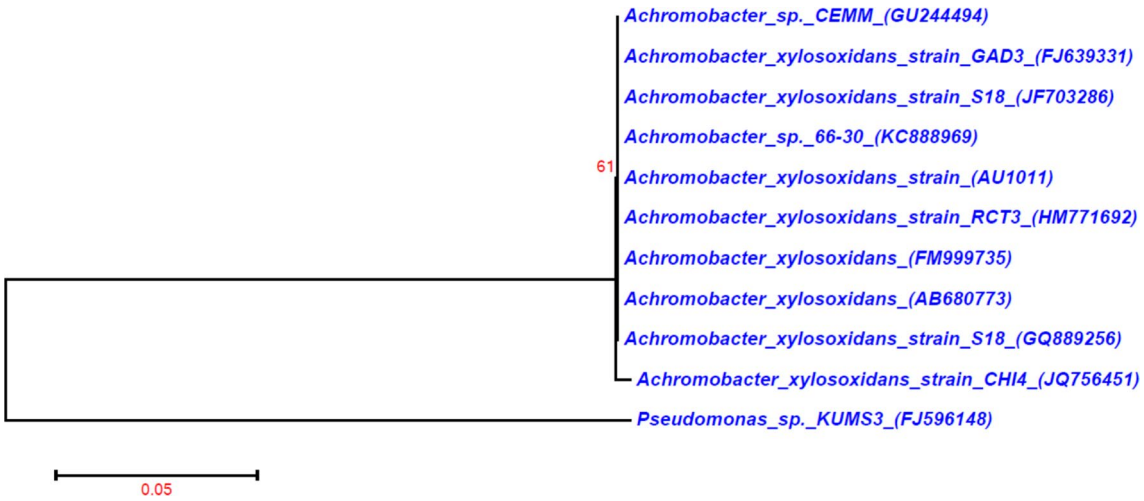


Figure 2. Phylogenetic tree of *A. xylosoxidans* with *Pseudomonas* sp. KUMS3 as outgroup.

	Protein (mg)	Total activity (U)	Specific activity (U/mg)	Purification fold	Recovery %
Culture supernatant	200	716	3.58	1.0	100
Ammonium sulphate precipitate	65	351	5.4	1.50	49
Sephadex column purified	8	86	10.75	3	12

Table 3. Estimation of *A. xylosoxidans* enzyme activity.

DXK012[32], *Pseudomonas aeruginosa* K-187[3], *Sphingomonas* sp. CJ-5[33], and *Serratia marcescens*[7]. A closely related study we were able to find was on an *Alcaligenes xylosoxidans* strain isolated from seafood industry waste[34,35].

Chitinase purification. The chitinase enzyme was effectively precipitated from the CHI4 culture supernatant system by salting-out method at 75% concentration of ammonium sulfate. After the dialysis and Sephadex-75 column chromatography purification, seven fractions containing enzyme activity were obtained. The enzyme specific activity increased to 10.75 U/mg after threefold purification using the Sephadex column, and the protein recovered was 8 mg (Table 3) and a total activity of 86 U/mg (Yield-12%) of chitinase enzyme were obtained. However, a high amount of recovery rate with 27% yield was noted in *Vibrio* sp.[36].

Effect of temperature and pH on the stability and activity of chitinase. The purified chitinase enzyme obtained from *A. xylosoxidans* was active 100% at an optimum temperature of 45 °C. The enzyme activity was observed at temperatures above 45 °C, but it decreased to 45% at 50 °C and lost its activity above 65 °C (Fig. 3). The optimum temperature for chitinase activity in various marine bacteria was reported in the past, i.e. 55 °C for *Bacillus cereus* SV1[28], 50 °C for *Alcaligenes xylosoxydans*[35] and *Alteromonas* sp. O-7[29], 40 °C for *Aeromonas hydrophila*[30], *Pseudoalteromonas* sp. DC14[31], *Pseudoalteromonas* sp. DXK012[32], *Pseudomonas aeruginosa* K-187[3] and 36 °C for *Sphingomonas* sp. CJ-5[33] , 30 °C for *Serratia marcescens*[7]. However, in the present study, the optimum temperature for chitinase activity ranged between 30–50 °C and was able to withstand up to 65 °C.

Chitinase from *A. xylosoxidans* was active in pH ranging between 7 and 9 and was found to have an optimum pH of 8.0. Moreover, the enzyme was 83% stable and 80% active at pH 8.0 (Fig. 4). The activity of chitinase reduced at pH beyond and below 6 due to the changes of ionic residues in the protein molecules. The ionic amino acid residue, which is found at the active site of the protein, is critical to maintaining protein conformation structure. Our results compared with the chitinase produced by the *Alteromonas* sp. O-7[29]. Also, the *Streptomyces chilikensis* RC1830 produced chitinase was active between pH 5–8 with an optimum pH 7[37]. The activity of chitinase produced from certain marine bacterias like *Sphingomonas* sp. CJ-5[33] and *Serratia marcescens*[7] has an optimum pH of 7, and the optimum pH was 9 for *Pseudoalteromonas* sp. DC14[31]. Whereas, chitinase of *Alcaligenes xylosoxydans* was active in acidic pH[35] and herein, *A. xylosoxidans* chitinase was active in a wide range of pH from neutral to base with 80% activity.

Screening of significant variables by Plackett–Burman design. Totally twelve variables were selected for analysing the effects on chitinase production by *A. xylosoxidans* using the Plackett–Burman design. The variables showing a positive impact on the enzyme production were selected for the optimisation study, and the significant variables of chitinase production with corresponding responses are exposed in Table 1. Besides, colloidal chitin, with an effect of 25.0, $MgSO_4$ (22.33) was the next significant factor to influence chitinase production, followed by peptone (17.53) and KH_2PO_4 (0.53), while the other variables, Yeast extract (− 1.40), $FeSO_4$

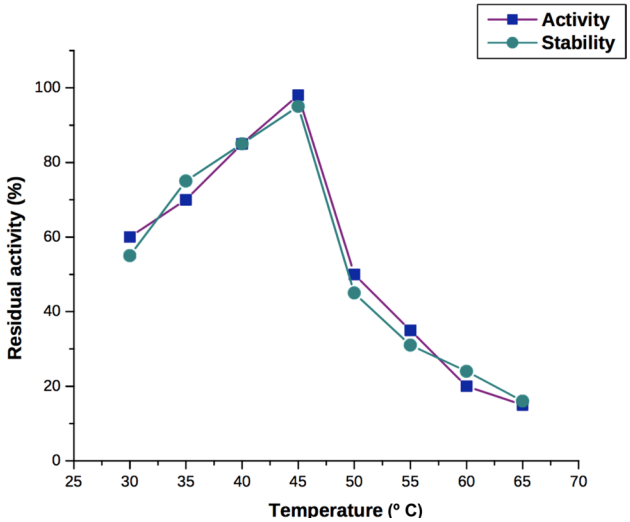


Figure 3. Effect of temperature on activity and stability of purified chitinase from *A. xylosoxidans.*

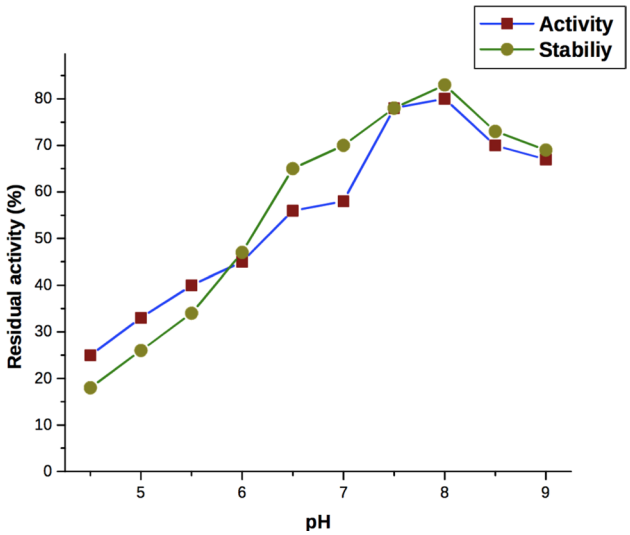


Figure 4. Effect of pH on activity and stability of purified chitinase from *A. xylosoxidans.*

(− 2.93) and $ZnSO_4$ (− 16.04) had a negative impact on the chitinase production (Fig. 5 and Table 4). Significant variables were identified, i.e., Colloidal chitin, peptone, KH_2PO_4 and $MgSO_4$ were selected along with the incubation period (37 °C for 3–5 days) to determine the optimum chitinase production by CCD further.

Similar experimental variables such as yeast extract, urea, KH_2PO_4 and colloidal chitin, were found to influence chitinase production by *Paenibacillus* sp.[38], however, in the present study peptone showed a positive impact. Colloidal chitin, syrup of date, yeast extract, K_2HPO_4 and KH_2PO_4 showed a positive impact on chitinase production in *Streptomyces griseorubens* C9[39]. The colloidal chitin as the sole carbon source, showed enhanced chitinase production in *Microbispora* sp.[40]. In contrast, better chitinase production was observed in the marine bacteria, *Aeromonas hydrophila* SBK1 when shrimp shell powder was used rather than colloidal chitin[41]. Phosphate was reported to have a negative effect on the chitinase production in *Streptomyces griseus*[42], but herein, it did have a modest positive effect on chitinase production (K_2HPO_4).

Peptone and yeast extract, the nitrogen sources (organic) did increase our chitinase production. Similar to the studies demonstrating chitinase production from *Alcaligenes xylosoxydans*[34], *Serratia marcescens*[43] and *Paenibacillus* sp. D1[38]. In contrast to this, *Paenibacillus* sp. AD showed maximum chitinase synthesis when ammonium sulfate was used as a nitrogen source[44], while sodium nitrate was ideal for the production of chitinase in *Stachybotrys elegans*[45]. In the present study, KH_2PO_4 and $MgSO_4$ showed a positive impact on chitinase production, whereas, their influence was insignificant during chitinase production by *Pseudomonas fluorescens* strain HN1205[46].

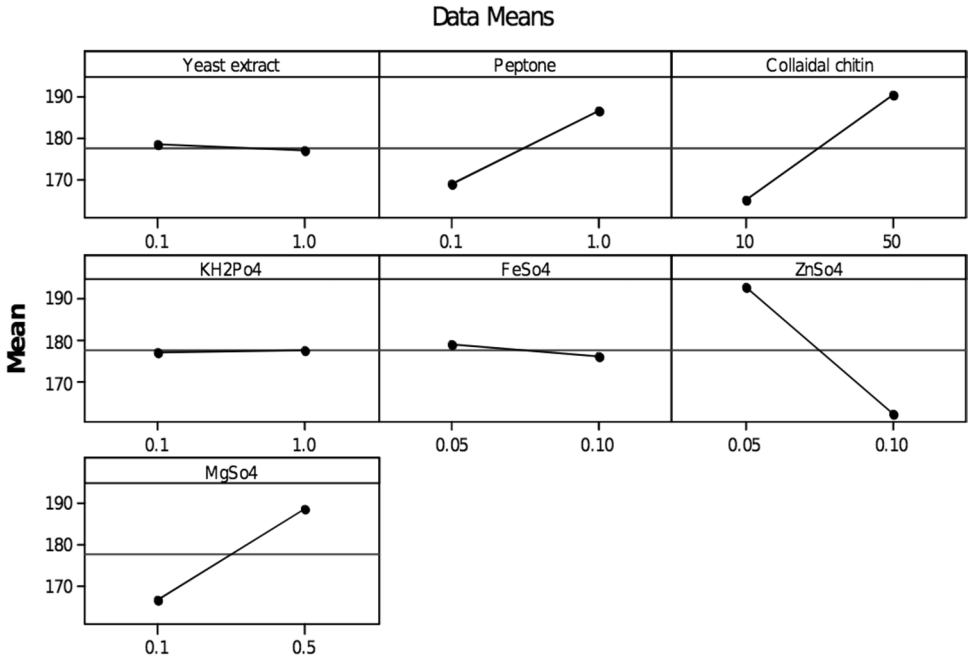


Figure 5. Main effects plot for chitinase production.

Term	Effect	Coefficient
Constant		177.63
Yeast extract	− 1.40	− 0.70
Peptone	17.53	8.77
Colloidal chitin	25.00	12.50
KH_2Po_4	0.53	0.27
$FeSo_4$	− 2.93	− 1.47
$ZnSo_4$	− 30.27	− 15.13
$MgSo_4$	22.33	11.17

Table 4. Result of the Plackett–Burman design based screening experiment.

Source	DF	Seq SS	Adj SS	Adj MS	F	*P*
Regression	14	45,770.8	45,770.8	3269.34	177.47	0.0001
Linear	4	5898.8	5898.8	1474.71	80.05	0.0001
Square	4	31,492.1	31,492.1	7873.02	427.37	0.0001
Interaction	6	8379.9	8379.9	1396.65	75.81	0.0001
Residual error	16	294.7	294.7	18.42		
Lack-of-fit	10	218.8	218.8	21.88	1.73	0.260
Pure error	6	76.0	76.0	12.67		
Total	30	46,065				

Table 5. Analysis of variance for chitinase production.

Optimisation of significant variables by CCD. The CCD experiments were conducted with four different variables (peptone, chitin, KH_2PO_4, $MgSO_4$) to determine their optimum concentration for the maximum production of chitinase. The ANOVA analysis for this model is presented in Table 5, and the *P* value for the model was 0.0001, *P* value of lack of fit was 0.260, suggesting that the data was a good fit with the model.

The three-dimensional response surface plots (regression equation in graphical representations) are shown in Figs. 6, 7, 8, 9, 10 and 11. The primary objective of the response surface is to ensure the variables' optimum values response is maximised powerfully. The surface plots affirm that the objective function was unimodal,

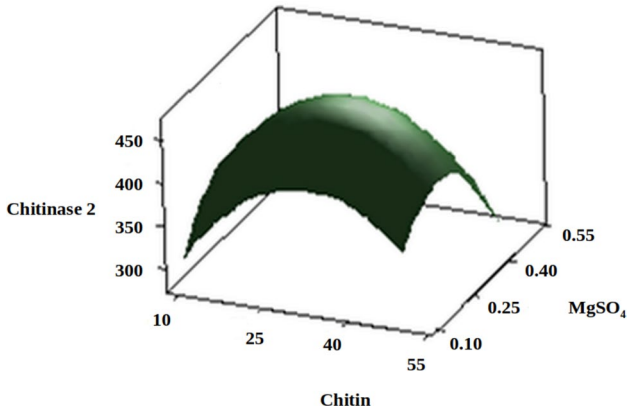


Figure 6. Response surface curve showing the effect of chitin and $MgSO_4$ on chitinase production by *A. xylosoxidans*.

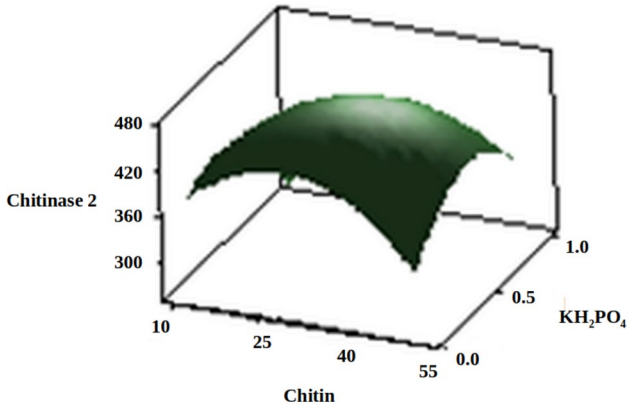


Figure 7. Response surface curve showing the effect of chitin and KH_2PO_4 on chitinase production by *A. xylosoxidans*.

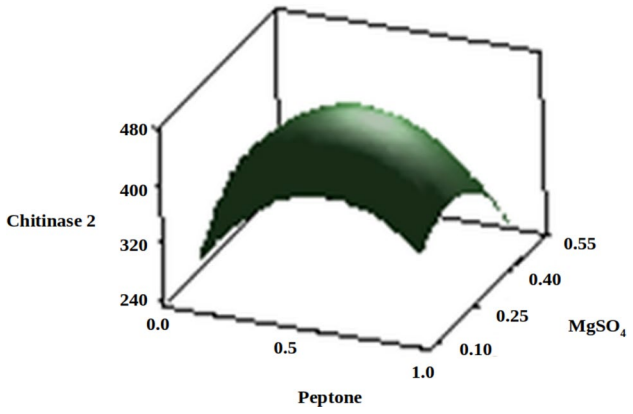


Figure 8. Response surface curve showing the effect of peptone and $MgSO_4$ on chitinase production by *A. xylosoxidans*.

which showed an optimum in the centre, and that the model is in a good prediction of the experiment. The main optimum point was estimated by the gradient method with the direction of the medium was the steepest rise for the chitinase production and estimated from the surface plots. The optimal values of peptone, chitin, KH_2PO_4 and $MgSO_4$ were estimated in actual units, and they were 0.5545 g/l, 30.6061 g/l, 0.50 g/l and 0.2818 g/l, respectively, with a chitinase activity of 464.2 U/ml (predicted). To confirm the predicted values, experiments

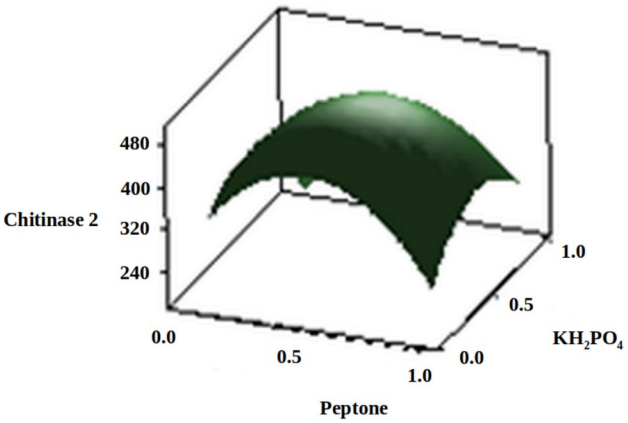


Figure 9. Response surface curve showing the effect of peptone and KH_2PO_4 on chitinase production by *A. xylosoxidans*.

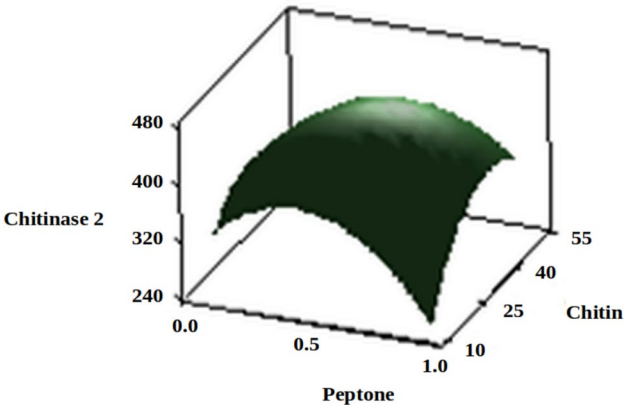


Figure 10. Response surface curve showing the effect of peptone and chitin on chitinase production by *A. xylosoxidans*.

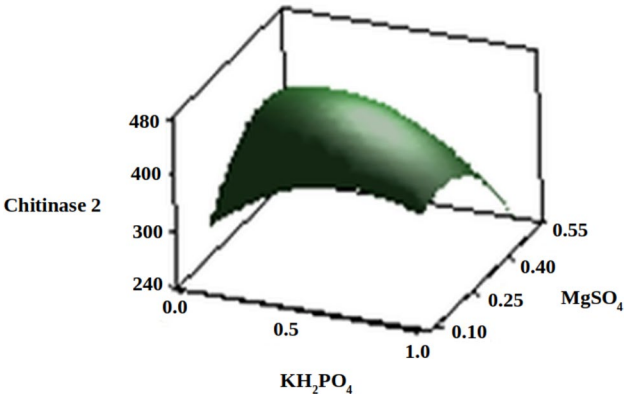


Figure 11. Response surface curve showing the effect of KH_2PO_4 and $MgSO_4$ on chitinase production by *A. xylosoxidans*.

were performed with these predicted conditions, and the recorded chitinase activity was 467 U/ml (Fig. 12). This activity was slightly high when compared to the predicted value. Such studies reiterate the fact that CCD is a highly reliable model.

Furthermore, the recorded chitinase activity was also higher than the chitinase production observed in most marine bacteria isolated from shrimp shell waste. A review on chitinase production and the produced quantity

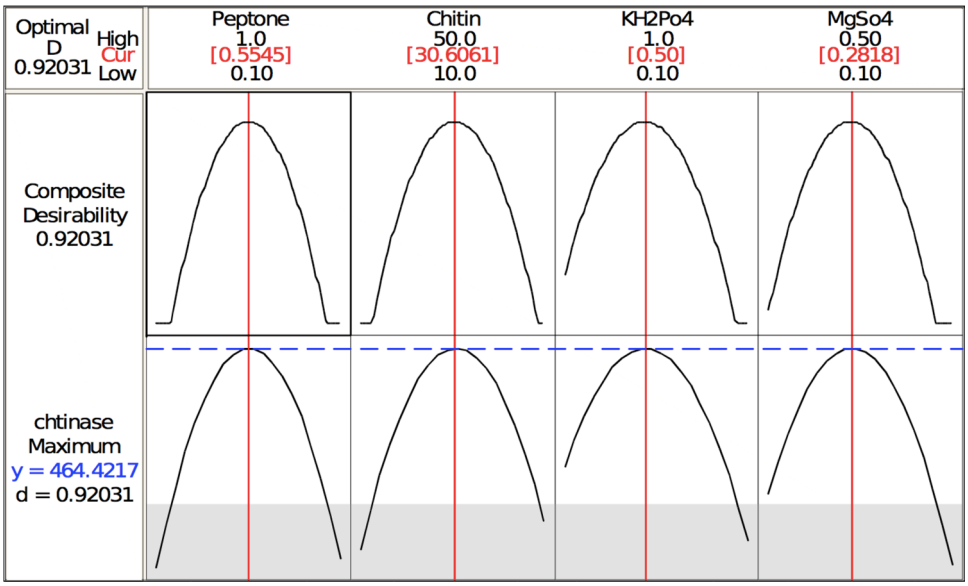


Figure 12. The predicted optimum values of peptone, chitin, KH_2PO_4 and $MgSO_4$ for chitinase production.

Species	Source of isolation	Quantity in U/ml	RSM-quantity in U/ml
Streptomyces sp.[49]	Shrimp shell waste	–	31.62
Bacillus cereus SV1[28]	Fishing port	–	82.8
Aeromonas hydrophila[41]	Shrimp shell waste	–	21.4 (BBD)
Pseudoalteromonas sp. DC14[31]	Caspian Sea	2.30	21.90 U/dl (BBD)
Bacillus pumilus[50]	Mangrove soil	3.36	23.19
Streptomyces griseorubens C9[39]	Semi-arid soil	–	9.7 (PBD and CCD)
Bacillus sp. CH-2[51]	Fish market soil	–	3.1 (PBD and CCD)
Paenibacillus sp.[44]	Coastal soil	3.84	20.01 (PBD and CCD)
Paenibacillus elgii[52]	Marine soil	3.157	24.53
Pseudomonas fluorescens strain HN1205[46]	Coastal soil	–	1.03 U/mL(PBD and CCD)
Paenibacillus sp. D1[38]		–	93.2 (PBD and CCD)

Table 6. Review on chitinase production and the recorded quantity by isolated marine bacterial strains.

by isolated marine bacterial strains can be seen in Table 6. Whereas, one should note that chitinase production from the marine fungi *Aspergillus flavus* MK20 isolated from coastal sediment can produce chitinase up to 620.5 U/ml, with RSM (Box-Behnken Design-BBD) based estimation up to 1520 U/ml[47].

Optimisation of *Paenibacillus* sp. D1, a marine microbial produced chitinase was demonstrated using CCD (predicted 85.8 U/ml) with an experimental design including 30 different formulations yielded 93.2 U/ml[38]. Whereas, herein, the maximum similarity was recorded between the experimental and the predicted results. Similar studies have also reported a positive effect of using statistical methods such as CCD to increase chitinase production by 141% in *Alcaligenes xylosoxidans*[35].

Amino acid analysis of chitin degraded product. The chitin degraded end products were analysed by HPLC. Usually, the chitin derivatives are oligomeric and monomeric polysaccharides, and few proteins are present. The results showed that in total, 15 amino acids were present (Fig. 13 and Table 7). The primary amino acid profile composition of the chitin degraded products includes lysine, aspartic acid, arginine, and low levels of isoleucine and methionine. Presence of major amino acids within the derivatives makes it an ideal aquaculture supplement.

While most prior studies focused on the amino acid profile of the chitinase enzyme, we have investigated the amino acid profile of chitin degraded products. Which lead to an exciting find that the amino acid profile of chitin degraded products has more composition along with traces of chitinase when compared with the amino acid profile of chitinase enzyme. Likewise, asparagine, threonine, glycine and alanine were reported as the major amino acids in chitin derivatives[48].

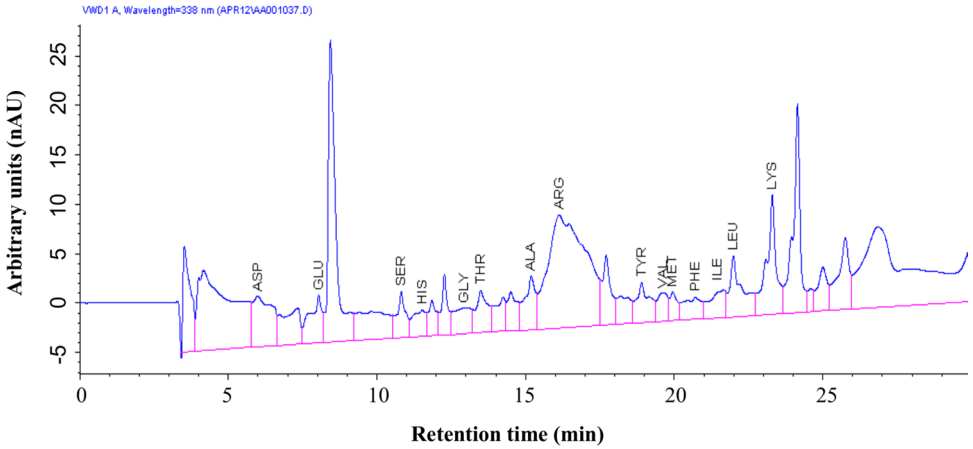


Figure 13. Amino acid composition of chitin degraded product by HPLC analysis.

Amino acids	Chitinase degraded product (ng/ml)
Aspartic acid	23.9
Glutamic acid	14.7
Serine	12.7
Histidine	9.2
Glycine	11.5
Threonine	15.6
Alanine	17.5
Arginine	61.3
Tyrosine	13.0
Valine	15.2
Methionine	10.9
Phenylalanine	11.0
Isoleucine	9.1
Leucine	19.7
Lysine	37.6

Table 7. Estimated amino acid composition of chitin degraded product.

Conclusion

In conclusion, we isolated 14 different strains of marine bacteria capable of degrading chitin. Out of that, the bacterial strain that had high chitinase activity was isolated. We have demonstrated that this *Achromobacter xylosoxidans* strain isolated from a shrimp waste disposal site can produce chitinase, which in turn can degrade chitin along with the concomitant production of N-acetyl-D-glucosamine. The simultaneous production of value-added bioproducts makes this an eco-friendly solution for environmental pollution caused by shrimp waste disposable. Moreover, these end products could also be promoted as a cost-effective supplement for the aquaculture industry. Furthermore, we have also demonstrated that this chitinase production by *Achromobacter xylosoxidans* can be optimised using RSM techniques.

Received: 8 January 2020; Accepted: 22 June 2020
Published online: 17 July 2020

References

1. Dhillon, G. S., Kaur, S., Brar, S. K. & Verma, M. Green synthesis approach: extraction of chitosan from fungus mycelia. *Crit. Rev. Biotechnol.* **33**, 379–403 (2012).
2. Mejia-Saules, J. E., Waliszewski, K. N., Garcia, M. A. & Cruz-Camarillo, R. The use of crude shrimp shell powder for chitinase production by *Serratia marcescens* WF. *Food Technol. Biotechnol.* **44**(1), 95–100 (2006).
3. Wang, S. L. & Chang, W. T. Purification and characterization of two bifunctional chitinases/lysozymes extracellularly produced by *Pseudomonas aeruginosa* K-187 in a shrimp and crab shell powder medium. *Appl. Environ. Microbiol.* **63**(2), 380–386 (1997).
4. Yadav, M. *et al.* Seafood waste: a source for preparation of commercially employable chitin/chitosan materials. *Bioresour. Bioprocess.* **6**(8), 1–20 (2019).
5. Das, S., Roy, D. & Sen, R. Chapter two—utilization of Chitinaceous wastes for the production of chitinase. In *Advances in food and nutrition research* (eds Kim, S.-K. & Toldrá, F.) 27–46 (Elsevier, Amsterdam, 2016). https://doi.org/10.1016/bs.afnr.2016.04.001.

6. Wang, S. L. *et al.* Reclamation of chitinous materials by bromelain for the preparation of antitumor and antifungal materials. *Bioresour. Technol.* **99**(10), 4386–4393 (2008).
7. Chakrabortty, S., Bhattacharya, S. & Das, A. Optimization of process parameters for chitinase production by a marine isolate of *Serratia marcescens*. *Int. J. Pharm. Biol. Sci.* **2**(2), 8–20 (2012).
8. Deeba, F., Abdullah, S. H., Irfan, M. & Iqbal, Q. J. Chitinase production in organisms: a review. *Punjab Univ. J. Zool.* **31**(1), 101–106 (2016).
9. Rathore, A. S. & Gupta, R. D. Chitinases from bacteria to human: properties, applications, and future perspectives. *Enzyme Res.* https://doi.org/10.1155/2015/791907 (2015).
10. Krithika, S. & Chellaram, C. Isolation, screening, and characterization of chitinase producing bacteria from marine wastes. *Int. J. Pharm. Pharm. Sci.* **8**(5), 34–36 (2016).
11. Tamadoni Jahromi, S. & Barzkar, N. Marine bacterial chitinase as sources of energy, eco-friendly agent, and industrial biocatalyst. *Int. J. Biol. Macromol.* **120**, 2147–2154 (2018).
12. Fondi, M. *et al.* "Every gene is everywhere but the environment selects": global geolocalization of gene sharing in environmental samples through network analysis. *Genome Biol. Evol.* **8**, 1388–1400 (2016).
13. Montgomery, D. C. *Design and analysis of experiments* 1–757 (Wiley, New York, 2008).
14. Huang, X., Wang, Y., Cui, Y. & Hua, X. Optimization of antifungal effect of surfactin and iturin to *Penicillium notatum* in syrup of peach by RSM. *Int. J. Pept. Res. Ther.* **16**, 63–69 (2010).
15. Treichel, H., Oliveira, D., Mazutti, M. A., Di Luccio, M. & Oliveira, J. V. A review on microbial lipases production. *Food Bioprocess Technol.* **3**, 182–196 (2010).
16. Yu, X., Hallet, S. G., Sheppard, J. & Watson, A. K. Application of the Plackett–Burman experimental design to evaluate nutritional requirements for the production of *Colletotrichum coccodes* spores. *Appl. Microbiol. Biotechnol.* **47**, 301–305 (1997).
17. Ergun, M. & Mutlu, S. F. Application of a statistical technique to the production of ethanol from sugar beet molasses by *Saccharomyces cerevisiae*. *Bioresour. Technol.* **73**, 251–255 (2000).
18. Plackett, R. L. & Burman, J. P. The design of optimum multifactorial experiments. *Biometrika* **37**, 305–325 (1946).
19. Simpson, T. W., Peplinski, J. D., Koch, P. N. & Allen, J. K. Metamodels of computer-based engineering design: survey and recommendations. *Eng. Comput.* **17**, 129–150 (2000).
20. Bergey, D. H. & Holt, J. G. *Bergey's manual of determinative bacteriology* (Lippincott Williams & Wilkins, Philadelphia, 2000).
21. Kumaran, S., Deivasigamani, B., Alagappan, K. M., Sakthivel, M. & Guru Prasad, S. Isolation and characterization of *Pseudomonas* sp. KUMS3 from Asian sea bass (*Lates calcarifer*) with fin rot. *World J. Microbiol. Biotechnol.* **26**(2), 359–363 (2010).
22. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98 (1999).
23. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**(12), 2725–2729. https://doi.org/10.1093/molbev/mst197 (2013).
24. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**, 265–275 (1951).
25. Reissig, J. L., Strominger, J. L. & Leloir, L. F. A modified colorimetric method for the estimation of N-acetylamino sugars. *J. Biol. Chem.* **217**, 959–966 (1955).
26. Yabuuchi, E. *et al.* Emendation of genus *Achromobacter* and *Achromobacter xylosoxidans* (Yabuuchi and Yano) and proposal of *Achromobacter ruhlandii* (Packer and Vishniac) comb. nov., Achromobacter piechaudii (Kiredjian et al.) comb. nov., and *Achromobacter xylosoxidans* subsp. denitrificans (Ruger and Tan) comb. nov. *Microbiol. Immunol.* **42**, 429–438 (1998).
27. Jenks, P. J. & Shaw, E. J. Recurrent septicaemia due to "*Achromobacter* group B". *J. Infect.* **34**, 143–145 (1997).
28. Ghorbel-Bellaaj, O., Manni, L., Jellouli, K., Hmidet, N. & Nasri, M. Optimization of protease and chitinase production by Bacillus cereus SV1 on shrimp shell waste using statistical experimental design. Biochemical and molecular characterization of the chitinase. *Ann. Microbiol.* **62**, 1255–1268 (2011).
29. Tsujibo, H. *et al.* Site-directed Mutagenesis of Chitinase from *Alteromonas* sp. strain O–7. *Biosci. Biotechnol. Biochem.* **57**(8), 1396–1397 (1993).
30. Hiraga, K. *et al.* Isolation and characterization of chitinase from a flake-chitin degrading marine bacterium, *Aeromonas hydrophila* H-2330. *Biosci. Biotechnol. Biochem.* **61**, 174–176 (1997).
31. Makhdoumi, A., Dehghani-Joybari, Z., Mashreghi, M., Jamialahmadi, K. & Asoodeh, A. A novel halo-alkali-tolerant and thermotolerant chitinase from *Pseudoalteromonas* sp. DC14 isolated from the Caspian Sea. *Int. J. Environ. Sci. Technol.* **12**(12), 3895–3904 (2015).
32. Liu, Y., Chan, Z. & Zeng, R. Purification and characterization of chitinase secreted by *Pseudoalteromonas* sp. DXK012 isolated from deepsea sediment. *J. Nat. Sci.* **1**, 1–5 (2015).
33. Zhu, X., Zhou, Y. & Feng, J. Analysis of both chitinase and chitosanase produced by *Sphingomonas* sp. CJ-5. *J. Zhejiang. Univ. Sci. B* **8**(11), 831–838 (2007).
34. Vaidya, R. J., Shah, I. M., Vyas, P. R. & Chhatpar, H. S. Production of chitinase and its optimization from a novel isolate *Alcaligenes xylosoxydans*: potential antifungal biocontrol. *World J. Microbiol. Biotechnol.* **17**, 62–69 (2001).
35. Vaidya, R. *et al.* Purification and characterization of chitinase from *Alcaligenes xylosoxydans*. *Biotechnol. Lett.* **25**, 715–717 (2003).
36. Park, S. H., Lee, J.-H. & Lee, H. K. Purification and characterization of chitinase from a marine bacterium, Vibrio sp. 98CJll027. *J. Microbiol.* **38**(4), 224–229 (2000).
37. Ray, L. *et al.* Purification and characterization of an extracellular thermo-alkali stable, metal tolerant chitinase from *Streptomyces chilikensis* RC1830 isolated from a brackish water lake sediment. *Biotechnol. Rep.* https://doi.org/10.1016/j.btre.2019.e00311 (2019).
38. Singh, A. K., Mehta, G. & Chhatpar, H. S. Optimization of medium constituents for improved chitinase production by *Paenibacillus* sp. D1 using statistical approach. *Lett. Appl. Microbiol.* **49**(6), 708–714. https://doi.org/10.1111/j.1472-765x.2009.02731.x (2009).
39. Meriem, G. & Mahmoud, K. Optimization of chitinase production by a new *Streptomyces griseorubens* C9 isolate using response surface methodology. *Ann. Microbiol.* **67**(2), 175–183 https://doi.org/10.1007/s13213-016-1249-8 (2006).
40. Nawani, N. N., Kapadnis, B. P., Das, A. D., Rao, A. S. & Mahajan, S. K. Purification and characterization of a thermophilic and acidophilic chitinase from *Microbispora sp*. V2. *J. Appl. Microbiol.* **93**, 965–975 (2002).
41. Halder, S. K. *et al.* Proficient biodegradation of shrimp shell waste by *Aeromonas hydrophila* SBK1 for the concomitant production of antifungal chitinase and antioxidant chitosaccharides. *Int. Biodeterior. Biodegrad.* **79**, 88–97 (2013).
42. Smucker, R. A. & Kim, C. K. Effects of phosphate on *Streptomyces griseus* chitinase production. In *Chitin, chitosan and related enzymes* (ed. Zikaki's, J.) 397–406 (Academic Press, New York, 1984).
43. Monreal, J. & Reese, E. T. The chitinase of *Serratia marcescens*. *Can. J. Microbiol.* **15**, 689–696 (1969).
44. Kumar, A., Kumar, D., George, N., Sharma, P. & Gupta, N. A process for complete biodegradation of shrimp waste by a novel marine isolate *Paenibacillus* sp. AD with simultaneous production of chitinase and chitin oligosaccharides. *Int. J. Biol. Macromol.* **109**, 263–272 (2018).
45. Tweddell, R. J., Jabaji-Hare, S. H. & Charest, P. M. Production of chitinases and β-1,3 glucanases by *Stachybotrys elegans*, a mycoparasite of *Rhizoctonia solani*. *Appl. Environ. Microbiol.* **60**, 489–495 (1994).
46. Lee, Y. S. & Kim, K. Y. Statistical optimization of medium components for chitinase production by *Pseudomonas fluorescens* strain HN1205: role of chitinase on egg hatching inhibition of root-knot nematode. *Biotechnol. Biotechnol. Equip.* **29**(3), 470–478 (2015).

47. Rawway, M., Beltagy, A. E., Abdul-Raouf, U. M., Elshenawy, M. A. & Kelany, M. S. Optimization of process parameters for chitinase production by a marine *Aspergillus Flavus* MK20. *J. Ecol. Health Environ.* **6**(1), 1–8 (2018).
48. Yong, T. *et al.* Purification and characterization of an extracellular chitinase produced by bacterium C4. *Ann. Microbiol.* **55**(3), 213–218 (2005).
49. Thiagarajan, V. R. *et al.* Extra cellular chitinase production by *Streptomyces* sp. PTK19 in submerged fermentation and its lytic activity on *Fusarium oxysporum* PTK2 cell wall. *Int. J. Curr. Sci.* **1**, 30–44 (2011).
50. Rishad, K. S., Rebello, S., Nathan, V. K., Shabanamol, S. & Jisha, M. S. Optimised production of chitinase from a novel mangrove isolate, *Bacillus pumilus* MCB-7 using response surface methodology. *Biocatal. Agric. Biotechnol.* **5**, 143–149 (2016).
51. Kumar, A. *et al.* A statistical approach for the chitinase production by *Bacillus sp*. and its antifungal potential. *IJESR* **5**(10), 28–47 (2017).
52. Philip, N. V. *et al.* Statistical optimization for coproduction of chitinase and beta 1, 4-endoglucanase by chitinolytic *Paenibacillus elgii* PB1 having antifungal activity. *Appl. Biochem. Biotechnol.* **191**, 135–150 (2020).

Acknowledgements

The study was funded by the CSIR (37(1425)10/EMR-11). Currently, MS is funded by the EPSRC and Imperial College London (EP/N509486/1: 1,979,819). We thank our funders.

Author contributions

Conceptualization, K.S.; Data curation, R.T. and S.B.; Formal analysis, W.A., B.S., R.T. and A.E.; Investigation, K.S.; Methodology, K.S. and A.K.; Project administration, D.B.; Resources, P.N., P.S.R. and A.E.; Software, B.S.; Supervision, D.B.; Validation, A.K. and P.N.; Visualization, W.A. and B.S.; Writing—original draft, K.S., B.S., G.P.S.; Data Visualization, Editing and Rewriting, M.S.

Competing interests

The authors declare no conflict of interest.

Additional information

Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020

ScienceDirect

Biocatalysis and Agricultural Biotechnology
Volume 29, October 2020, 101802

Bioactive metabolites produced from *Streptomyces enissocaesilis* SSASC10 against fish pathogens

Subramanian Kumaran [a, b, 1], selvaraj Bharathi [c, 1], Vairagkar Uttra [f], Rajasekar Thirunavukkarasu [b], Prasannabalaji Nainangu [c], Venugopal Gopi Krishnan [b], Pugazhvendan Sampath Renuga [e], Aruni Wilson [a], Deivasigamani Balaraman [d]

[a] School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology, Tamil Nadu, India
[b] Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Tamil Nadu, India
[c] Department of Microbiology, Sri Sankara Arts and Science College, Tamil Nadu, India
[d] Centre for Advanced Studies in Marine Biology, Annamalai University, Tamil Nadu, India
[e] Department of Zoology, Annamalai University, Tamil Nadu, India
[f] Savitribai Phule Pune University, Pune, 41100, Maharashtra, India

Received 10 July 2020, Revised 17 September 2020, Accepted 23 September 2020, Available online 1 October 2020, Version of Record 10 October 2020.

Check for updates

Show less

Share Cite

https://doi.org/10.1016/j.bcab.2020.101802 Get rights and content

Abstract

In this study Marine pigment producing actinobacteria were screened for bioactive metabolite against fish pathogens. Totally ten morphologically different actinobacteria was isolated from sediment sample of Mudasalodai, Chidambaram, India. The isolates were preliminary tested for antibacterial activity by agar plug method against Pseudomonas sp, Vibrio anguillarum, Aeromonas hydrophilia, *Vibrio paraheamolyticus* and Vibrio harveyi. The brown pigment producing isolate SSASC found to inhibit all the tested fish pathogens and hence selected for further study. The bioactive metabolite was extracted using different solvents such as methanol, chloroform, ethyl acetate and n-hexane in which ethyl extract showed promising antibacterial activity. The minimum inhibitory concentration (MIC) of the ethyl acetate extract ranged from 15.6 μg/ml to 125 μg/ml and minimum bactericidal concentration (MBC) found to be in the range of 62.5 μg/ml to 250 μg/ml. The ethyl acetate extract was purified using TLC and a single spot with Rf value of 0.45 was observed in TLC. The FT-IR spectrum exhibits characteristic signals at 16 peaks corresponding to functional groups indicating phosphenes, amino acids, and presence of alkyl ether. GC-MS analysis of brown pigment from SSASC showed 25 peaks. The potent strain SSASC10 was identified as Streptomyces *enissocaesilis* based on morphological, biochemical and molecular characteristics. In this present finding, the isolated *Streptomyces enissocaesilis* SSASC10 strain found to have excellent antibacterial activity against fish pathogens hence can be used for treating infectious fish disease.

Introduction

Aquaculture sector has reached tremendous growth in the last decade. Many different strategies have to be developed to further improvise the profitability (FAO, 2012). The main threat in aquaculture industry because of infectious diseases results in monetary

loss to aquaculturists. The challenge in the cultivation of fish and shellfish is prevention of bacterial disease (Bachere et al., 1995). Various environmental factors (oxygen levels, pH, temperature, water quality, osmotic strength, availability of iron) and poor management practices such as over feeding, inadequate nutrition, overcrowding results in lower growth rate and increase in susceptibility of disease outbreaks (Verschuere et al., 2000). Bacterial infections including *Vibrio harveyi, V. parahaemolyticus, V. campbellii* producing luminous vibriosis in aquaculture industry results in 5–100% mortality of shrimp (Wang et al., 2015). In addition to *Vibrio* sp., diverse species of *Pseudomonas* also cause opportunistic infections in fish, producing great economic losses in aquaculture industry. Another opportunistic common disease causing bacterial pathogen is *Aeromonas hydrophila* causing infections in warm water fishes and also in cold water fishes. Bacterial diseases are often requiring treatment with antibiotics added to feeds or water (Idowu et al., 2017). The uncontrolled use of antibiotics results in emergence of antibiotic resistant bacteria that pose a major threat in aquaculture industry. The ideal solution to this problem is the development of alternative antibiotics for disease control and treatment.

Actinobacteria considered as intermediate group between bacteria and fungi has obtained a distinct position (Pandey et al., 2004). Actinobacteria are prokaryotes that were isolated from various natural and man-made ecosystems (Debananda et al., 2009). They have high guanine and cytosine in their DNA that constitutes up to 55%. Actinobacteria are recognized as rich source of metabolites such as antibiotics, antitumor agents, immune - suppressive agent and enzymes hence, considered as economically and biotechnologically valuable prokaryotes.

Streptomyces are the dominant group among the actinobacteria that are widely distributed in nature. *Streptomyces* are the potential candidate for the production of antibiotics and other industrially important compounds (Dharmaraj et al., 2009). Some of the examples include aminoglycosides, ansamycines, aminocyclitoles, β-lactams, chloramphenicols, cyclic peptides, glycopeptides, lincosamides, tetracyclines, macrolides, streptogramines, phenylpropanoids (Sing et al., 2000). The prolific source of antibiotic producer are actinobacteria in particular *Streptomyces* sp since it provides more than half of the naturally occurring antibiotics and still needs to be explored further for novel antibiotics (Berdy, 2005). Marine actinobacteria possess diverse and superior characteristics than terrestrial actinobacteria in production of secondary metabolites (Blunt et al., 2007). The present study aims in the exploitation of marine actinobacteria against fish bacterial pathogens.

Section snippets

Sample collection

Soil sample was collected from Mudasalodai (Lat 11° 29′ N; long.79°46′ E) landing centre, shrimp and crab disposable area, along the Parangipettai coast situated at South East coast of India, Tamil Nadu....

Isolation of pigmented actinobacteria

For isolation of actinobacteria, the collected soil sample was serially diluted and plated on starch casein agar (Starch-1 g, Casein-0.03 g, KNO_3-0.2 g, $MgSO_4$-0.005 g, $FeSO_4$-0.001 g, $CaCO_3$-0.002 g, NaCl-0.2 g, Agar-1.5 g per 100 ml) and incubated at 28 °C for 5–7 days. After incubation,...

Isolation of actinobacteria

The actinobacterial strains were isolated from the site of shrimp waste disposable areas, Parangipettai, Cuddalore district. Morphologically 10 different strains of actinobacteria were isolated using starch casein agar as the isolation medium. Priyanka et al. (2019) reported totally isolated nine morphologically different actinobacteria using starch casein agar from the salterns of Ongole, Andhra Pradesh....

Preliminary screening of actinobacteria against fish borne pathogens

Bacterial infections are considered the major cause of mortality in aquaculture. They can...

Conclusion

In conclusion, this study describes the isolation of pigmented actinobacteria from marine sediments. The isolated actinobacteria showed promising antibacterial activity against fish pathogens. The bioactive metabolite was extracted and characterized by FTIR and GC-MS. The potent actinobacteria was identified as *Streptomyces enissocaesilis* SSASC10. Furthermore investigations are needed for the possible use of bioactive metabolite of *Streptomyces enissocaesilis* SSASC10 for possible use in...

CRediT authorship contribution statement

selvaraj Bharathi: Conceptualization, Data curation, Visualization, Formal analysis, Investigation, Methodology, Resources, Software, Writing - original draft, Writing - review & editing. **Vairagkar Uttra:** Conceptualization, Data curation, Visualization, Methodology, Resources, Software. **Rajasekar Thirunavukkarasu:** Conceptualization, Data curation, Visualization, Methodology, Resources, Software. **Prasannabalaji Nainangu:** Conceptualization, Data curation, Visualization, Methodology, Resources,...

Acknowledgement

Authors all are great full thanks to Dr. S.T. Balasubramanian, Former Dean, CAS in Marine Biology, Parangipettai for giving lab facilities. The Authors express their heart full thanks to Management of Sathyabama Institute of Science and Technology, Chennai, for research facilities. Authors are grateful thanks to Prof. K.R. Venkatesan, Principal, Sri Sankara Arts and Science College for his constant support....

References (29)

S. Priyanka *et al.*
Characterisation and identification of antibacterial compound from marine actinobacteria: In vitro and in silico analysis
Journal of Infection and Public Health (2019)

A. Kavitha *et al.*
Purification and biological evaluation of the metabolites produced by *Streptomyces sp.* TK-VL-333
Res. Microbiol. (2010)

E. Bachere *et al.*
Knowledge and research prospects in marine mollusc and crustacean immunology
Aquaculture (1995)

H.M. Atta *et al.*
Sparsomycin antibiotic production by *Streptomyces sp.* AZ-NIOFD1: taxonomy, fermentation, purification and biological activities
Am. Eur. J. Agric. Environ. Sci. (2009)

J. Berdy
Bioactive microbial metabolites: A personal view
J. Antibiot. (2005)

J.W. Blunt *et al.*
Marine natural products
Nat. Prod. Rep. (2007)

S.N. Debananda *et al.*
Screening of actinomycete isolates from niche habitats in Manipur for antibiotic activity
Am. J. Biochem. Biotechnol. (2009)

S. Dharmaraj
Marine Streptomyces as a novel source of bioactive substances
World J. Microbiol. Biotechnol. (2010)

S. Dharmaraj *et al.*
Fermentative production of carotenoids from marine Actinomycetes
Iran. J. Microbiol. (2009)

N.E. El-Naggar *et al.*

In vitro* activity, extraction, separation and structure elucidation of antibiotic produced by *Streptomyces anulatus* NEAE-94 active against multidrug-resistant *Staphylococcus aureus

Biotechnol. Biotechnol. Equip. (2017)

View more references

Cited by (5)

Antimicrobial and antioxidant activities of Streptomyces species from soils of three different cold sites in the Fez-Meknes region Morocco

2022, Scientific Reports

Antibacterial Activity of Carotenoid Pigments Produced by Heterotrophic Bacteria from Seawater in Krakal Coastal Area, Yogyakarta, Indonesia

2022, Squalen Bulletin of Marine and Fisheries Postharvest and Biotechnology

Actinomycetes as a promising candidate bacterial group for the health management of aquaculture systems: A review

2022, Reviews in Aquaculture

Enzyme production, antibacterial and antifungal activities of actinobacteria isolated from Ficus deltoidea rhizosphere

2022, Biodiversitas

Incidence, Antibiotic Susceptibility and Characterization of Vibrio parahaemolyticus Isolated from Seafood in Selangor, Malaysia

2021, Progress in Microbes and Molecular Biology

Recommended articles (6)

Research article

In vitro biomedicinal properties of Pyrrolidine-2,4-Dione derived from a novel actinobacterium *Streptomyces rochei*, a green approach

Biocatalysis and Agricultural Biotechnology, Volume 20, 2019, Article 101244

Show abstract

Research article

Bioactive compounds of *Streptomyces*: Biosynthesis to applications

Studies in Natural Products Chemistry, Volume 64, 2020, pp. 467-491

Show abstract

Research article

Systematic characterization of potential cellulolytic marine actinobacteria *Actinoalloteichus* sp. MHA15

Biotechnology Reports, Volume 13, 2017, pp. 30-36

Show abstract

Research article

Extraction and partial purification of secondary metabolites from endophytic actinomycetes of marine green algae *Caulerpa racemosa* against multi drug resistant uropathogens

Biocatalysis and Agricultural Biotechnology, Volume 17, 2019, pp. 750-757

Show abstract

Research article

Isolation, structure elucidation and antibacterial activity of methyl-4,8-dimethylundecanate from the marine actinobacterium *Streptomyces albogriseolus* ECR64

Microbial Pathogenesis, Volume 121, 2018, pp. 166-172

Show abstract

Research article

Melanin from marine *Streptomyces* sp. (MVCS13) with potential effect against ornamental fish pathogens of *Carassius auratus* (Linnaeus, 1758)

Biocatalysis and Agricultural Biotechnology, Volume 3, Issue 4, 2014, pp. 134-141

Show abstract

[1] Equal contribution.

View full text

© 2020 Elsevier Ltd. All rights reserved.

Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9316120

Automating Guide selection process of the Department through Web Application

T. Sasikala [1], S.Vigneshwari [2*], BandreddiSandeep[3]

[1,2,3] School of Computing, Sathyabama Institute of Science and Technology Chennai,Tamilnadu, India.
([1]sasi_madhu2k2@yahoo.co.in, [2]vikiraju@gmail.com,, [3]bandreddisandeep@gmail.com)

ABSTRACT

This is a website where guide selection was implemented. In this process every student has the access to view the profile of all teachers to register them as their project guides. First come First serve procedure is followed in this site. Selection of guides is based on the specialization of the Faculties of the Department. The main agenda is to assign faculties strong at particular areas to students' projects based on that area. Faculty name has to be registered with all relevant details. The student register his/her details in the portal. The faculty register his/her details in the portal. The student checks for availability of the guide. If the guide is available then the student selects the guide and the guide slot availability is decremented by one. If the guide is not available the student has to select another guide from the portal. The list of selected students can be viewed by faculty. The guide selection detail will be displayed on the student portal. The complete details can be downloaded as a CSV file from the admin login. Technologies used for implementation are: HTML with bootstrap, Node Js/PhP, MySql database, Cloud Server: It is published through amazon web services.

Keywords : Guide selection, web application, registration, Bootstrap, PhP, Java Script

1. INTRODUCTION

Guide selection process is developed in PHP without using any frameworks. In this website Bootstrap version 4 is used with Html and Css for front-end designing, PhP is used as backend and sql as the database for the site. Google authentication is used for authenticating users to access the website. Tabs are very easy to access and Compatible with both mobile and desktop views independent on browsers.

As Google authentication is utilized here, it is more secured so that all other third party logins are avoided. By this security students can feel comfortable while using this application without any obligations. This comprises very quick and easy steps to select a teacher based on their specialisation. User friendly interfaces which helps end users with more comfort. Very quick responses are received as a strong backend and effective database are utilized.

Normally, project Guides are assigned to all students in a random procedure which have a drawback when a student with a machine learning project is assigned to a guide who has his specialisation in Internet of Things. To overcome this kind of drawbacks Guide selection process application helps a lot of students as well as Guides. The final outcome of this application benefits guides and students by allotting students having projects based on their specialization.

Guide Selection process is based on First come First Serve basis for students and to participate in this process guides need to be registered in a secured link where all their details will be recorded. The order of displaying guides for students is based on their ID number. This ID number is used as an unique id for guides to participate in this process. Maximum 10 students can register for each guide and can modify based on the total strength. The server used in this application is very efficient and allows a lot of students at a time without the server crashing.

The application gives an excel sheet of batch wise list for every guide individually which also include project name and the partners name. Alternatively registered students can see their guide in the guide selection website and it does not allow them to register again. An individual student can register to guide either as a team or as an individual, Team registration requires the partner's registration number and his name. This whole process saves a lot of time for administration to assign guides to every student for their projects. Simple signup, login and select a guide based on their specialization whether it is related to your project and Logout.

2. RELATED WORK

Pooja et al[15] developed a framework for College Automation for K.C. College of Engineering & Management Studies & Research – Thane (E), India and has said it in production successfully. This gives us a motivation for development of a similar system, the automated guide selection system which is considered as a part of college automation.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:15:37 UTC from IEEE Xplore. Restrictions apply.

Subbiah et al[14] developed an automated set up for measuring student performance and record maintenance system for centralize administration which reduces the time consumption of manual maintenance.

Bharamagoudar et al[16] developed a web based ERP for automating student data and Faculty data. The system forms a centralized database for maintaining and analysing the entire institutions- student faculty data with respect to library, placement, accounts department and so on.

Vigneshwari et al[13] proposed a Machine learning approach which is applicable for any real time automation systems. Sai et al[12] developed another institutional system which helps in recognizing and alerting the management in the arrival of strangers.

3.WORKING MODEL

In this web application there are seven interfaces totally. Basically these interfaces are divided into parts, one for admins and other for students who are the users of this application. Admin pages are secured and only accessible by admins whereas student accessible pages are public and are available to all users registered.

Login interface and Signup interface are the important interfaces in accessing this application, only students with a valid registration number can register for this application. Guide selection process fully based on first come first serve basis especially for students who want to select their project. Every guide will have a particular limit to get allotted and if this limit exceeds the guide cannot be selected further to any of the students. Registration process works differently for both students and guides, student registration needs a valid registration number and an email whereas a guide needs an Id number given to them by their college and an email. If a user forgets his/her password they can reset their password in the forgot password section by providing their email and date of birth which are given during registration. Once a registered student cannot register again for a guide, so just register for the best guide who will be more helpful in your project.

After Login interfaces users can see a page with all guides brief details like name, designation, specialisation and experience, any user can view their profile by clicking on the view button where he can see more details like subjects taught, projects completed and email which will be more helpful in contacting a guide. A student can register either as a team or as an individual in this interface. If a student is already registered then he will see the name of that particular teacher on top of the screen. New teachers can register to this application by admin login, during this registration they need to provide their unique id given to them by the university.

Admin logins have access to every page in the application. Admin can see the students registered for each teacher by entering the teacher id, he can also review the details of all teachers. If any teacher wants to update his/her profile they need to contact the adminto edit their details. Every teacher will have a limit and once the limit is reached students cannot register to that particular teacher no more. Once a registered student cannot change their guid or their project partner. To change their project guide they need to get permission from the registered guide and their head of the department.

This paper studies the guide selection process to students for their project submission. This is a web application to select the project guides who will be more helpful to students during their projects. In this application every student can register using their gmail and register number which is more secured during. Instead of allotting guides to students in a random pool which involves a lot of process this online application can decrease the amount of work to be done and save time.

The application is handled in two phases, one is the front end part and the backend part. Both of these phases are studied detail in the following

- Front end Phase – HTML
- Back end Phrase– PhP
- Database - Sql

3.1. HTML

Html is used as the front end in this application for documents design. It is also assisted by CSS(Cascading Style Sheets) and Javascript. Online servers have html documents and they are sent to web browsers. Structure of the web page is described by Html. Functions can be written in scripting language like javascript to affect behaviour and content of web pages. The look and Layout are defined by adding CSS in Html Files. Tags are used in these files which are enclosed in angle brackets like <h1>. Tags can have other tags inside them which refers to nested tags or mixture of tags. These tags also have attributes which indicate identifiers. Start tags are identified as and end tags as . Some tags like
 don't have any end tags . These require only a single empty tag which breaks the line by using this tag. An anchor element in html makes a hyperlink in the document and the href attribute inside the anchor tag sets the target URL which links to another document. To render an image in HTML, img tag is used along with image URL as the attribute in the tag. Css can also be declared inline or it can be linked with a link tag that refers to the file where cssof a particular document can be found. Html also provides some tags like <input> to take inputs from users which are mostly used form tags. Id attribute provides a document wide unique identifier. This identifier is used in stylesheets to change its properties and they are also used in scripts for animation and designing processes[8,9,10,11].

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:15:37 UTC from IEEE Xplore. Restrictions apply.

3.2. Bootstrap

Bootstrap is a front-end framework which is entirely based on cssstylings and scripts. This is also a veryresponsive framework which is designed both for mobile devices and desktops. Bootstrap comprises many pre-design templates for forms , buttons, navigation and many other interfaces based on javascript. This web framework gives basic style definitions for almost all Html elements. Many JQueryplugins are also included for several components which provide other additional user interface elements. The basic component to place all bootstrap elements is container and this container is of two types namely fixed-width container and a fluid-width container. Container makes all elements inside it to implement a CSS Flex box layout by rows and columns.

3.3. Java Script

Javascript is also one of the core technologies like Html and CSS of the World Wide Web and an essential part of web applications. Majority of websites use it for client-side page behaviour and it is also a multi-paradigm language. It supports a lot of styles like event-driven, functional and imperative programming styles. Javascript has API(application programming interfaces) for working with text, dates, regular expressions, DOM(Document Object Model) and standard data structures. This is the dominant client-side language of the web and used for every animation in the guide selection process like loading new page content without loading the whole page, posting to a database also done by Ajax without leaving the page. Animation like fading them in and out, resizing and moving them. Side Panels in this website are animated through jquery[5,6,7]

.

3.4.PhP

PhP is used as a backend in this application to connect with a database. All operations like fetching, storing, deleting and updating databases are done using PhP. It is a popular general-purpose scripting language especially for web development,PhP can only be processed in web servers. On a server, the result of executed PhP code like generated Html or binary data would form the whole or part of a Http response. PhP code can also run through a local server or by command line interface. Sql database is used to store the data of students, teachers and students responses and this database is connected to a server via PhPmysqli connection. PhP also includes various free and open-source libraries in resulting PhP binary builds. PhP mainly used for dynamic web page content, this provides session variables to check stop breaking other non registered users from accessing this website. All validations are done based on registrations and these validations are implemented using PhP code by comparing them to stored data and giving particular users certain limits[1,2,3,4].

Figure 1 depicts the overall workflow of the model.

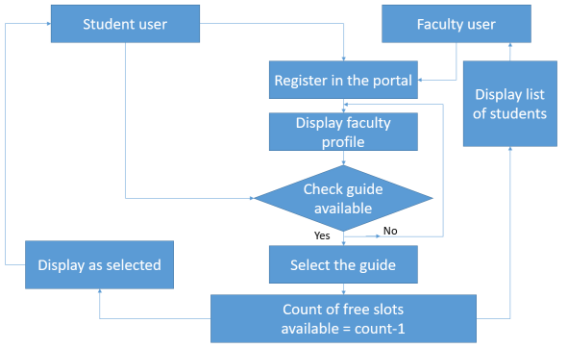


Figure 1 Overall workflow

4. Results and Discussions

The implementation is done using Html with the help of Bootstrap and PhP by using a cloud server. It is also designed to load the total leaves and their reasons. MySQL database is created for storing the data of the students, student registrations and total data of the faculty registered.

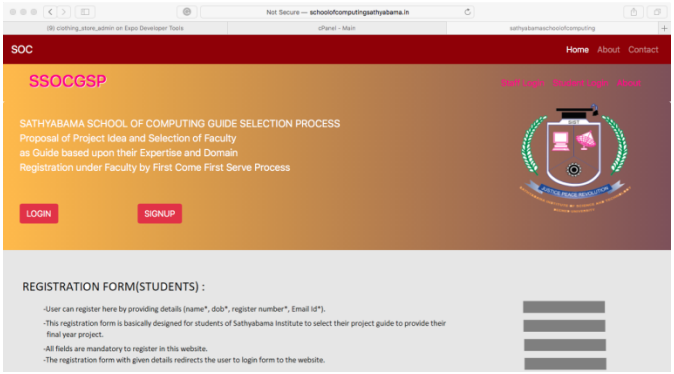


Figure 2. a. Home page

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:15:37 UTC from IEEE Xplore. Restrictions apply.

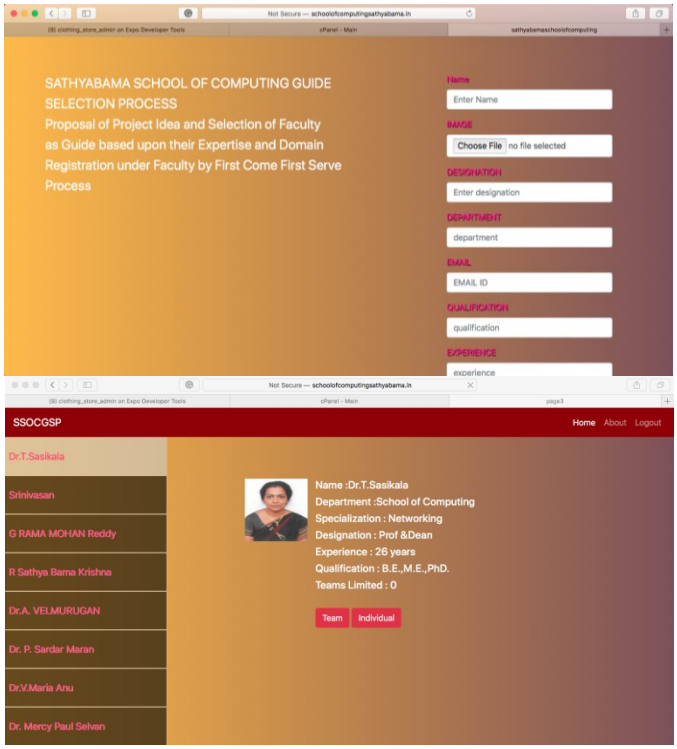


Figure 2. b. Faculty Registration



Figure 2. c. Student guide selection form

Figure 2. a. depicts the Home page of the portal developed. There are separated logins for students and faculties. Before a student selects a faculty, the faculty has to register with the portal along with photograph and area of specialization details(Figure 2.b) Once the faculty registration is over, the students can select their project guide faculty whose area of research maps the students area of interest. Figure 2.c shows the student guide selection form. A maximum limit of students who can select a particular faculty has been allotted.

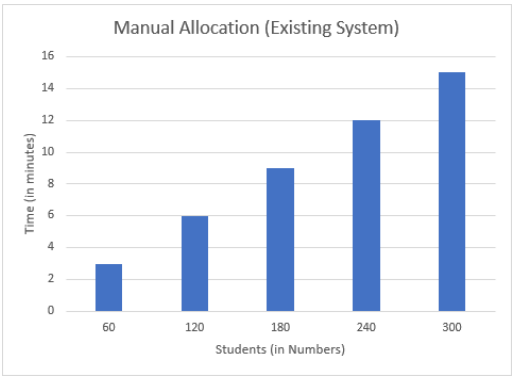


Figure 3. Existing system performance

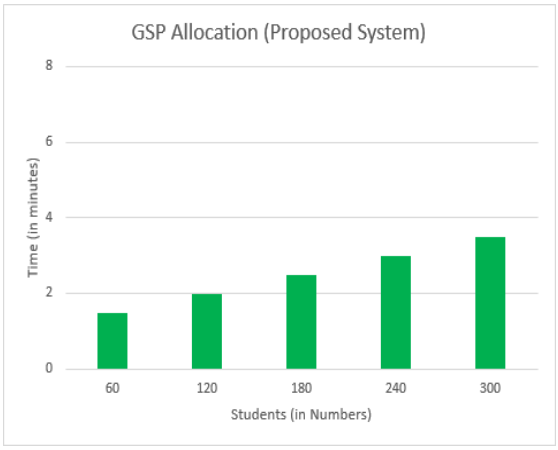


Figure 4. Proposed system efficiency

Figure 3 shows the comparison chart with the existing guide selection or the student information system. Figure 4 shows our proposed Guide Selection Portal(GSP) which shows improved performance in units of time for the same selected batches of students.

5. CONCLUSION

Guide selection process was entirely made of PhP latest version, this has many advantages and can be used for many years and it is also more efficient as it is published through cloud services. This web application decreases the time consumed by faculty in allotting students to random guides

REFERENCES

[1] Basic Coding Standard. PHP Framework Interoperability Group, https://www.php-fig.org/psr/psr-1/ ,[online] Last Accessed, May 2020.

[2] "Description of core php.ini directives – Manual. php.net. ,https://www.php.net/manual/en/ini.core.php#ini.short-open-tag,[online],Last Accessed, May 2020.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:15:37 UTC from IEEE Xplore. Restrictions apply.

[3] Problems with PHP. http://toykeeper.net/soapbox/php_problems/,[online] Last Accessed, May 2020.

[4] "php.net 2007 news archive". The PHP Grouphttps://www.php.net/archive/2007.php . ,[online] Last Accessed, May 2020.

[5] "JavaServer Pages Technology — JavaServer Pages Comparing Methods for Server-Side Dynamic Content White Paper". Sun Microsystems. http://178.32.81.228/products/jsp/jspservlet.html ,[online] Last Accessed, May 2020.

[6] "JavaScript data types and data structures - JavaScript | MDN". Developer.mozilla.org. https://developer.mozilla.org/en-US/docs/Web/JavaScript/Data_structures ,[online] Last Accessed, May 2020.

[7] "Usage statistics of JavaScript as client-side programming language on websites". w3techs.com. https://w3techs.com/technologies/details/cp-javascript ,[online] Last Accessed, May 2020.

[8] "HTML 5.1 is a W3C Recommendation". World Wide Web Consortium. 1 November 2016. https://www.w3.org/blog/news/archives/5932,[online] Last Accessed, May 2020.

[9] "HTML5". World Wide Web Consortium. June 10, 2008. https://en.wikipedia.org/wiki/World_Wide_Web_Consortium, ,[online] Last Accessed, May 2020.

[10] W3Schools: https://www.w3schools.com, ,[online] Last Accessed, May 2020.

[11] "HTML: The Markup Language (an HTML language reference)". https://www.w3.org/TR/2012/WD-html-markup-20121025/ ,[online] Last Accessed, May 2020.

[12] Sai, M. Yashwanth, R. Vijai Chandra Prasad, P. R. Niveditha, T. Sasipraba, S. Vigneshwari, and S. Gowri. "Low cost automated facial recognition system." In 2017 *Second International Conference on Electrical, Computer and Communication Technologies (ICECCT),* pp. 1-5. IEEE, 2017.

[13] Vigneshwari, S., Bharathi, B., Sasikala, T., Mukkamala, S.A study on the application of machine learning algorithms using R,Journal of Computational and Theoretical Nanoscience,2019,*Journal of Computational and Theoretical Nanoscience,*16(8), pp. 3466-3472

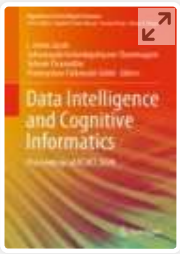
[14] Dr. K. Venkata Subbiah, D. Dinesh, Ch Suresh,'Development of a Student Database Management System for a University, *Int. Journal of Engineering Research and Application* , Vol. 6, Issue 8, (Part -2) 2016, pp.16-24

[15] Pooja S. Sharma,Reshma R. Shetty,Gayatri V. Yadkikar,Dhanashri Kanade, 'College Automation System',International *Journal for Innovative Research in Science & Technology,* Vol 2, Issu 10, 2016, pp. 96-102

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:15:37 UTC from IEEE Xplore. Restrictions apply.

Springer Link

Search Log in



Conference proceedings | © 2021

Data Intelligence and Cognitive Informatics

Proceedings of ICDICI 2020

Editors: I. Jeena Jacob, Selvanayaki Kolandapalayam Shanmugam, Selwyn Piramuthu, Przemyslaw Falkowski-Gilski

Presents research works in the field of data intelligence

Discusses results of ICDICI 2020 held in Tirunelveli, India, during 8–9 July 2020

Serves as a reference for researchers and practitioners in academia and industry

Part of the book series: Algorithms for Intelligent Systems (AIS)

50k Accesses | **148** Citations

Sections

Table of contents

About this book

Keywords
Editors and Affiliations
About the editors
Bibliographic Information

This is a preview of subscription content, access via your institution.

Table of contents (74 papers)

Search within book

← Previous Page 1 of 4 Next →

Front Matter PDF
Pages i-xxii

Text Generation and Enhanced Evaluation of Metric for Machine Translation
Sujit S. Amin, Lata Ragha
Pages 1-17

Destress It—Detection and Analysis of Stress Levels
Neha Udeshi, Nemil Shah, Urvi Shah, Stevina Correia
Pages 19-33

On the Design and Applications of an Integrated Smart Home Automation
Thanaz Shafeer, M. Senthil Arumugam, A. Sasikala
Pages 35-52

Voice and Gesture Based App for Blind People

Tilak Satra, Manali Shah, Ajay Lad, Stevina Correia
Pages 53-75

Digital Image Forgery Detection Using Ternary Pattern and ELM

D. Vaishnavi, D. Mahalakshmi, M. S. Shawon Rahman
Pages 77-86

A Study on Identification and Cleaning of Struck-Out Words in Handwritten Documents

R. Dheemanth Urs, H. K. Chethan
Pages 87-95

RFID-Based Railway Crowd Prediction and Revenue Analysis

Shraddha S. More, Gulshan Pathak, Simran Panchal, Monika Patil
Pages 97-109

A Review on Efficient Scheduling Techniques for Cloud Computing

Venkateswara Rao Kumpati, Manish Pandey
Pages 111-120

A Synonym-Based Bi-LSTM Model for Machine Reading Comprehension

S. Anbukkarasi, N. Abinaya, S. Varadhaganapathy
Pages 121-129

Designing of Arbiter PUF for Securing IP and IoT Devices

Swati Kulkarni, R. M. Vani, P. V. Hunagund
Pages 131-138

Enhancements on the Efficacy of Firefighting Robots Through Team Allocation and Path-Planning

Sreesruthi Ramasubramanian, Senthil Arumugam Muthukumaraswamy, A. Sasikala
Pages 139-151

Smart Meter Using Raspberry Pi for Efficient Energy Utilization

Chidrupa Manogna Mamunooru, Suseendra Attada, Sai Vipanchi Kodali, Mohan Reddy Tumma, Santanu Kumar Dash, Y. Padma Sai
Pages 153-163

An Unsupervised Content-Based Article Recommendation System Using Natural Language Processing

S. Renuka, G. S. S. Raj Kiran, Palakodeti Rohit
Pages 165-180

A Survey of the Exemplary Practices in Network Operations and Management

K. M. Majidha Fathima
Pages 181-194

Utilization of Fish Excrete for Plant Growth Using SVM

C. Gnana Kousalya, T. U. Gowdhami, R. Mayavinothini
Pages 195-206

A Novel Approach of Resource Scheduling Algorithm to Improve QoS in Green Cloud

Computing

P. Geetha, C. R. Rene Robin
Pages 207-221

Machine Learning Approach for Analysis of the Company Performance Based on Fundamental Data

Pratik Rathi, Madhur Kabra, Rahil Sheth, Jignesh Sisodia
Pages 223-231

Evaluation of Deep Learning Models in the Prediction of Lung Disease(Pneumonia)

Adusumilli Rohit, B. Padmaja, K. Vinay Kumar, T. Chandana, M. Madhu Bala
Pages 233-241

An Automated Invoice Handling Method Using OCR

Pranay Kumar, S. Revathy
Pages 243-254

← Previous Page 1 of 4 Next →

Back to top ↑

About this book

This book discusses new cognitive informatics tools, algorithms and methods that mimic the mechanisms of the human brain which lead to an impending revolution in understating a large amount of data generated by various smart applications. The book is a collection of peer-reviewed best selected research papers presented at the International Conference on

Data Intelligence and Cognitive Informatics (ICDICI 2020), organized by SCAD College of Engineering and Technology, Tirunelveli, India, during 8–9 July 2020. The book includes novel work in data intelligence domain which combines with the increasing efforts of artificial intelligence, machine learning, deep learning and cognitive science to study and develop a deeper understanding of the information processing systems.

Back to top ↑

Keywords

Intelligent Information Systems

Cognitive Intelligence

Industrial Internet of Things

Cognitive Human-Computer Interaction Models

Mobile Technologies and Services

ICDICI 2020 **ICDICI Proceedings**

Deep Learning

Back to top ↑

Editors and Affiliations

GITAM University, Bangalore, India
I. Jeena Jacob

Department of Mathematics and Computer Science, Concordia University Chicago, River Forest, USA
Selvanayaki Kolandapalayam Shanmugam

University of Florida, Gainesville, USA
Selwyn Piramuthu

Gdańsk University of Technology, Gdańsk, Poland
Przemyslaw Falkowski-Gilski

Back to top ↑

About the editors

I. Jeena Jacob is working as a Professor in Computer Science and Engineering department at GITAM University, Bangalore, India. She actively participates on the development of the research field by conducting international conferences, workshops and seminars. She has published many articles in refereed journals. She has guest edited an issue for the International Journal of Mobile Learning and Organisation. Her research interests include mobile learning and computing.

Selvanayaki Kolandapalayam Shanmugam is currently working as an Associate Professor in Computer Science, Concordia University, Chicago, USA. She had over all 15+ years of lecturing sessions for theoretical subjects, experimental and instructional procedure for laboratory subjects. She presented more research articles in the national & international conferences and journals. Her research interest includes image processing, video processing, soft computing techniques, intelligent computing, web application development, object-oriented programming like C++, and Java, scripting

languages like VBscript and Javascript, data science, algorithms, data warehousing and data mining, neural networks, genetic algorithms, software engineering, software project management software quality assurance, enterprise resource planning, information systems and database management systems.

Selwyn Piramuthu is Professor of Information Systems at the University of Florida. He received his B.Tech., M.S. and Ph.D., respectively, from IIT Madras, University of Arizona and University of Illinois at Urbana–Champaign. His research interests include machine learning and cryptography with applications in medical informatics, supply chain management, financial credit risk scoring, IoT, among others. His book, co-authored with Wei Zhou titled, "RFID and Sensor Network Automation in the Food Industry," was published by Wiley in 2016.

Przemyslaw Falkowski-Gilski is a graduate of the Faculty of ETI, Gdansk University of Technology. He graduated 1st degree B.Sc. studies (in Polish) and 2nd degree M.Sc. studies (in English) in 2012 and 2013, respectively. He pursued his Ph.D. studies in the field of electronic media, particularly digital broadcasting systems and quality of networks and services. In 2018, he received the title of Doctor of Technical Sciences with distinction, discipline Telecommunications, specialty in Radio communication. Currently, he works as an academic.

Back to top ↑

Bibliographic Information

Book Title	Book Subtitle	Editors
Data Intelligence and Cognitive Informatics	Proceedings of ICDICI 2020	I. Jeena Jacob, Selvanayaki Kolandapalayam Shanmugam,

Selwyn Piramuthu, Przemyslaw Falkowski-Gilski

Series Title	DOI	Publisher
Algorithms for Intelligent Systems	https://doi.org/10.1007/978-981-15-8530-2	Springer Singapore

eBook Packages	Copyright Information	Hardcover ISBN
Intelligent Technologies and Robotics, Intelligent Technologies and Robotics (R0)	The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021	978-981-15-8529-6 Published: 09 January 2021

Softcover ISBN	eBook ISBN	Series ISSN
978-981-15-8532-6 Published: 09 January 2022	978-981-15-8530-2 Published: 08 January 2021	2524-7565

Series E-ISSN	Edition Number	Number of Pages
2524-7573	1	XXII, 941

Number of Illustrations	Topics
149 b/w illustrations, 363 illustrations in colour	Computational Intelligence, Mathematical Models of Cognitive Processes and Neural Networks, Artificial Intelligence

Back to top ↑

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Artificial Cells, Nanomedicine, and Biotechnology
An International Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ianb20

Pseudomonas aeruginosa synthesized silver nanoparticles inhibit cell proliferation and induce ROS mediated apoptosis in thyroid cancer cell line (TPC1)

Jinmei Yang, Qiang Wang, Chunxiao Wang, Ruiyi Yang, Mukhtar Ahmed, Subramanian Kumaran, Periyannan Velu & Bo Li

To cite this article: Jinmei Yang, Qiang Wang, Chunxiao Wang, Ruiyi Yang, Mukhtar Ahmed, Subramanian Kumaran, Periyannan Velu & Bo Li (2020) *Pseudomonas aeruginosa* synthesized silver nanoparticles inhibit cell proliferation and induce ROS mediated apoptosis in thyroid cancer cell line (TPC1), Artificial Cells, Nanomedicine, and Biotechnology, 48:1, 800-809, DOI: 10.1080/21691401.2019.1687495

To link to this article: https://doi.org/10.1080/21691401.2019.1687495

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 20 May 2020.

Submit your article to this journal

Article views: 2247

View related articles

View Crossmark data

Citing articles: 13 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ianb20

ARTIFICIAL CELLS, NANOMEDICINE, AND BIOTECHNOLOGY
2019, VOL. 48, NO. 1, 800–809
https://doi.org/10.1080/21691401.2019.1687495

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Pseudomonas aeruginosa synthesized silver nanoparticles inhibit cell proliferation and induce ROS mediated apoptosis in thyroid cancer cell line (TPC1)

Jinmei Yang[a], Qiang Wang[b], Chunxiao Wang[c], Ruiyi Yang[a], Mukhtar Ahmed[d], Subramanian Kumaran[e], Periyannan Velu[f] and Bo Li[c]

[a]Physical Examination Center, The First Affiliated Hospital of Kunming Medical University, Kunming, China; [b]Physical Examination Center, The Fourth Affiliated Hospital of Kunming Medical University, Kunming, China; [c]Department of General Surgery, The Fourth Affiliated Hospital of Kunming Medical University, Kunming, China; [d]Department of Zoology, College of Science, King Saud University, Riyadh, Kingdom of Saudi Arabia; [e]Department of Biotechnology, School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology (Deemed to Be University), Chennai, India; [f]Scigen Research and Innovation Pvt. Ltd, Periyar Technology Business Incubator, Thanjavur, India

ABSTRACT

We used cell-free culture filtrate of *Pseudomonas aeruginosa* as a reducing mediator of $AgNO_3$ to silver-nanoparticles (AgNPs) and possibly used as a potential anticancer agent against thyroid cancer cells (TPC1). The bio-generation of AgNPs was firmly established by taking a UV spectrum at 380–500 nm wavelength. The Fourier transform spectrum analysis reveals the association of alcohol, phenol and aromatic functional groups with *P. aeruginosa* synthesized AgNPs (*Ps*-AgNPs). By observing under transmission electron microscopy (TEM), the size and structure of the *Ps*-AgNPs were characterized as the size was 30–70 nm and spherical in shape. The concentration-dependent cytotoxicity of *Ps*-AgNPs on TPC1 cells was observed and IC50 value was calculated. The alteration of oxidative and antioxidant biomarkers in Ps-AgNPs treated cells were observed. The induced apoptosis was determined by staining the *Ps*-AgNPs treated cells with DCFH-DA, Rh-123 dye, Acridine Orange (AO) and ethidium bromide (EtBr). Increased level of intracellular reactive oxygen species (ROS) generation and decreased level of mitochondrial membrane potential was observed in *Ps*-AgNPs treated TPC1 cells. Moreover, the apoptotic morphological changes were explored, which indicates increased apoptosis by inducing cell membrane damage in *Ps*-AgNPs treated cells. This biogenic approach will enable an effective and significant improvement in nano-oncotherapy.

ARTICLE HISTORY
Received 20 May 2019
Revised 7 August 2019
Accepted 10 August 2019

KEYWORDS
Silver nanoparticles; thyroid cancer; TPC1 cells; lipid peroxidation; nano-oncotherapy

Introduction

Nanoparticles synthesis is an important aspect in nanotechnology that has attained growing curious in nano-medicine technology as a result of their potential application in chemotherapies. AgNPs are broadly used in numerous fields like health care, biomedical, food and industrial purposes [1,2]. Nanotechnology tools provide an effective approach against cancers to overcome biological barriers to directly deliver therapeutic drugs [3]. The unique physicochemical properties metallic nanoparticles such as broad optical properties, high surface–volume ratio, ease of synthesis and surface functionalization provide new insights in chemotherapeutics.

Usually, the processes involved in AgNPs synthesis has detailed so far related to the high energy consumption and involvement of hazardous toxic chemicals. The pure biological molecules (like peptide and enzymes) aided AgNPs synthesis has been used successfully around the world [4–6]. This biomolecule approach in nanotechnology has a greater attraction in the field of nanoparticle aided chemotherapies. However, the green reactants approach for AgNPs synthesis based on algal and plant extracts significantly gained great interest due to its biocompatibility, eco-friendliness and benignancy. These appealing features are essential in medical applications. The other advantages of green synthesis include the well-defined and controlled size of the nanoparticles. They are devoid of contaminants and the process is easy to scale-up [7]. Moreover, microorganisms including bacteria, yeast and fungi are being swiftly subjugated as nano-factories in generating numerous nanomaterials [8,9].

The synthesis of nanomaterial by means of the microbial process provides an extensive variety of green and environmentally desirable technique; low-fee production and much less time required [10]. Therefore, the yield effective, harmless, consistent, clean and environmentally friendly methods have required for the preparation of the nanoparticles. The

CONTACT Bo Li lb18669179389@sina.com Department of General Surgery, The Fourth Affiliated Hospital of Kunming Medical University, Kunming 650021, Yunnan, China

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

microorganism for the production of silver nanoparticles (AgNPs) displays great interest in the formation of AgNPs because of its resistant potential to the metals. Synthesis of nanoparticles using cell filtrate could be efficient than the intracellular synthesis [11]. The microbial growth medium contains peptone, yeast extract, dextrose and other important growth factors, some of which may have potential reducing and stabilizing ability as stated in previous reports [12].

For AgNP synthesis, the microbial culture media was aspirated along with remaining microbial biomass after the complete growth of microorganisms. Either with or without the microbial biomass, the silver salt was added into the aspirated growth media. A specific set of enzymes(s) that are secreted by particular microorganisms are consequential elements to define the biosynthesis of nanoparticles, but the main factors for the formation of nanoparticles are not yet clear [13]. Several anticancer drugs were ineffectual to reach their goal due to insufficient concentrations and capably employ the pharmacological consequence without giving harmfulness to healthy tissues and cells [14,15]. Biosynthesized AgNPs, either it was generated by intra- or extracellular mechanisms have been effectively examined as a result of its potent antimicrobial ability against various microbial pathogens [16–18] and cytotoxic nature against various cancer cells [19].

Thyroid cancer is the most common endocrine malignancies among the other cancers. The incidence rate of thyroid cancer has stably increased over the past decade, which considered as one of the causes of death [20]. Papillary thyroid cancer (PTC) is the major histological type which accounts for approximately 80% of human thyroid cancers [21]. The prognosis of PTC has been proven to rely on several well-established clinicopathologic indicators [22]. In this study, we assessed the anticancer potential of *Pseudomonas aeruginosa* generated AgNPs against thyroid cancer cells (TPC1) and then revealed its oxidative stress-induced apoptosis.

Materials and methods

Biosynthesis of AgNPs

For biosynthesis of AgNPs, *Pseudomonas aeruginosa* bacterial strain was periodically cultured and maintained in Nutrient Agar (NA) slants. From 48 h old *P. aeruginosa* culture, one loop full of inoculum was thoroughly suspended in 5 ml of double-distilled water then the suspension was transferred directly to flask containing 45 ml of MGYP (malt extract 3 g, glucose 10 g, yeast extract 3 g and peptone 5 g in 1 l of distilled water and pH 7) broth medium, then the cells were subsequently grown for 48 h in orbital shaking incubator (200 rpm) at 27 °C. For AgNPs synthesis, the cell pellet was harvested by centrifuging the culture broth for 5 min at 6000 rpm and then eroded using 50 mM phosphate buffer (pH 7). The obtained cell pellet was thoroughly suspended in double distilled water then incubated in an orbital shaker at 200 rpm for 24 h. After incubation, the cell-free supernatant was obtained by centrifuging at 7500 rpm for 10 min. Finally, 5 ml of 1 mM $AgNO_3$ solution was transferred to 100 ml flask having 50 ml of collected cell-free culture supernatant, and then the conical flask was effectively cultured at 27 °C in an orbital shaker at 200 rpm for 20 min [23].

Evaluation of maximum AgNPs synthesis

In order to evaluate the maximum synthesis of nanoparticle at different incubation period (1 h, 24 h, 7 days and 15 days), totally four reactions (cell-free culture supernatant and $AgNO_3$) were prepared as described above. The absorption spectral readings of reactants were recorded by ultra violet–visible (UV–vis) spectrophotometer with 0.72 nm resolution (Shimadzu, Kyoto, Japan, UV Pharma spec 1700) from 300 to 800 nm.

Characterization of biosynthesized Ps-AgNPs

The size and structural features of the biosynthesized *P. aeruginosa* synthesized AgNPs (*Ps*-AgNPs) were characterized by transmission electron microscopy (TEM, JEOL-JEM-2010, Tokyo, Japan). The crystalline nature of the *Ps*-AgNPs was examined by X-ray diffraction analyzer (D/Max 2005, Rigaku, Tokyo, Japan). In order to explore the surface functional groups associated with synthesized AgNPs, the *Ps*-AgNPs subjected to Fourier transform infrared spectrophotometer analysis (FT-IR; Shimadzu 8201PC, Kyoto, Japan).

Cancer cell culture

The thyroid cancer cells (TPC1) were inoculated in 10% foetal bovine serum and 1% antibiotics (penicillin, streptomycin, and amphotericin) supplied Dulbecco's modified Eagle's medium (DMEM), incubated at 37 °C in a moistened atmosphere provided with 5% CO_2 and 95% air incubation. *Ps*-AgNPs were completely dissolved in 1% DMSO prior treatment with cell.

Cytotoxicity of Ps-AgNPs against TPC1 cells

The cytotoxicity of *Ps*-AgNPs was evaluated by employing 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay based on the reduction of MTT by the action of mitochondrial reductase enzyme which presents in viable cells. The amount of reduction of MTT to formazan in *Ps*-AgNPs treated and untreated cells were directly proportional to the number of viable cells, it was assessed by taking absorbance values by dissolving the formazan with DMSO. In brief, 1×10^5 cells were plated into 96-well plates and incubated for 24 h at 37 °C in a humidified condition. After complete incubation, the old medium was aspirated without disturbing cells then the cells were treated with different concentrations (5, 10, 20, 30, 40, 50, 60, 70 and 80 μg/ml) of *Ps*-AgNPs dissolved DMEM, the plate was kept in same incubation condition. After incubation, 100 μl of MTT reagent (5 mg/ml in PBS) was subsequently added into each well further, the plates were incubated for 24 h at 37 °C. The resulting formazan was dissolved by adding 100 μl of DMSO and the absorbance of the reactant solution was recorded at 595 nm wavelength using a multiwell plate reader (Tecan

Multimode Reader, Grödig, Austria). The concentrations of the test sample which showed 50% of cell death were then calculated [23].

Biochemical assays

The TPC1 cells were harvested after treatment of biosynthesized *Ps*-AgNPs (25, 50 and 75 μg) and subsequently subjected to biochemical assessments. The levels of lipid peroxidation in *Ps*-AgNPs treated cells were estimated by measuring LPO (lipid peroxide) byproduct of TBARS (thiobarbituric acid) reactive substance by the method described by Okhawa et al. [24], and the activities of enzymatic antioxidants such as superoxide dismutase (SOD) [25], catalase (CAT) [26], the intracellular enzyme of reduced glutathione (GPx) level [27] were measured. The untreated well was employed as a control for all assays, respectively.

Measurement of intracellular ROS generation

Intracellular reactive oxygen species (ROS) production in *Ps*-AgNPs treated and untreated TPC1 cells were estimated by non-fluorescent probe DCFH-DA that can easily enter into the intracellular cell matrix, there it was oxidized into fluorescent dichlorofluorescein (DCF) by the action of produced ROS. Accordingly, the fluorescence strength was relatively proportional to the level of ROS production [28]. The TPC1 cells were inoculated (1×10^6 cells/well) into six-well plate, treated with different concentrations (25, 50 and 75 μg) of *Ps*-AgNPs and placed in CO_2 (5%) incubator for 24 h. After incubation, cells were exposed to 100 μl of DCFH-DA for 10 min at 37 °C. Fluorescence depth was estimated with excitation and emission filters set at 485 and 530 nm, respectively (Shimadzu RF-5301 PC spectrofluorometer, Kyoto, Japan). The results showed an increased percentage of fluorescence depth.

Determination of mitochondrial membrane potential

The mitochondrial membrane potential (MMP) of treated and untreated TPC1 cells was estimated by staining the cells with Rhodamine-123 (Rh-123), lipophilic cationic dye [29]. The cells (3×10^4 cells/well) were cultured in six wells plate and treated with various concentrations of *Ps*-AgNPs (25, 50 and 75 μg) for 24 h. Then, Rh-123 fluorescent dye was added to treated and untreated cells and left for 30 min. The MMP was qualitatively examined under a FLoid cell imaging station (Invitrogen, Carlsbad, CA, USA). Subsequently, the cells were harvested by trypsinization of cells and the fluorescence strength was estimated at 485/530 nm wavelength under spectrofluorometer (Shimadzu, Columbia, MD, USA). The results were compared with positive control which maintained without treatment.

Determination of apoptotic morphological changes

Staining the treated and untreated TPC1 cells with Acridine Orange (AO) and ethidium bromide (EtBr) dye was used to determine the apoptotic cells morphological changes by *Ps*-AgNPs [30]. The cells were cultivated in six-well plate (3×10^4 cells/well) and treated with *Ps*-AgNPs at various dosage ranges (25, 50 and 75 μg) for 24 h. The cells were fixed in methanol:glacial acetic acid (3:1) for 30 min at 4 °C. The cells were stained with AO/EtBr (1:1 ratio) and left for 30 min after washing cells with PBS. The stained cells were consequently washed again with PBS and observed under a FLoid cell imaging station (Invitrogen, Carlsbad, CA, USA). The apoptosis was assessed by counting the cells showing apoptosis features in the total number of cells observed in the microscopic field.

Statistical analyses

All experiments were conducted in three independent replicated and the results were stated as the mean ± standard deviation (mean ± SD) analyzing by one-way analysis of variance (ANOVA) between the obtained values. Values of $p < .05$ indicated the significant differences in data.

Results

Biosynthesis of AgNPs

The biogenic generation of AgNPs by the cell-free extract of *P. aeruginosa* was initially confirmed by visual observation as shown in Figure 1(A). It was noticed that the no colour changes in $AgNO_3$ solution alone, which indicates that no convention of AgNPs. The colour change from pale yellow to brown was noticed in a flask containing the culture filtrate and $AgNO_3$, which denoted that the reduction of $AgNO_3$ to AgNPs. The colour change to dark brown was observed in extend incubation period up to 6 h.

Further, the *Ps*-AgNPs synthesis was investigated by taking spectral readings by UV–vis spectrophotometer at 380–500 nm wavelength. Figure 1(B) shows the UV–visible readings of AgNPs synthesized at various time intervals (1 h, 24 h, 7 days and 15 days) after added silver nitrate to cell filtrate resulted in the increasing spectrum of absorption at around 420 nm. The absorption spectrum of *Ps*-AgNPs displayed a greater wide peak at 420 nm and allocated to surface plasmon resonance of the *Ps*-AgNPs. Moreover, the increased spectrum at 420 nm was observed an extended incubation period.

Characterization of Ps-AgNPs

TEM

A TEM examination provides a better understanding of the size and structural features of the *Ps*-AgNPs. The TEM images at various magnifications ranges are depicted in Figure 2(A), which shows the shape and size of the bio-generated *Ps*-AgNPs. It revealed that the *Ps*-AgNPs sphere-shaped and the size were ranged from 30 to 75 nm. As shown in Figure 2(B), the average size of synthesized *Ps*-AgNPs was 70 mm and has 15.1% of occurrence.

Fourier transform infra-red (FTIR) spectroscopy analyses

In order to analyze the numerous functional groups accountable for the reduction of silver nitrate to AgNPs, the FTIR spectrum patterns were taken and subsequently interpreted. As shown in Figure 2(C), the spectrum depicts the number of peaks that were located at 3244, 2921, 2090, 1646, 1541 and 1009 cm^{-1}. The band arising at 3244 cm^{-1} in the spectrum resembles O–H stretching representing the occurrence of alcohol and phenol groups. A band at 2921 cm^{-1} region was assigned for C–H stretching of an aromatic compound. The peak at 2090 cm^{-1} region was arising from N=C=S stretching of isothiocyanate. A peak raised at 1646 cm^{-1} resembles C=C stretching of the conjugated alkene. A band at 1541 cm^{-1} assigned for strong N–O stretching of the nitro compound and the band at 1009 cm^{-1} strong C–F stretching of fluoro compound.

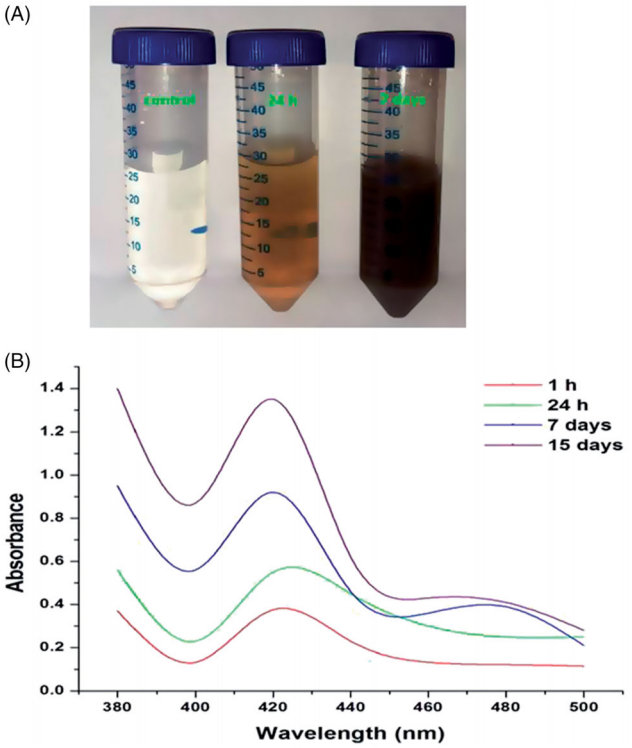


Figure 1. (A) Visual observations of colour change from fine yellow to reddish brown; indicates the reduction of $AgNO_3$ into silver nanoparticles. (B) UV–visible spectrum of biosynthesized silver nanoparticle at different time intervals (1 h, 24 h, 7 days and 15 days) at 380–500 nm. The increased spectrum at 420 nm explores synthesize of silver nanoparticle.

X-ray diffraction analyses

Figure 3 shows the XRD analysis established the crystal-like nature of *Ps*-AgNPs, that depicts peaks at 2θ values of four distinct diffraction peaks at 33.28°, 46.38°, 67.54° and 77.64° corresponding to XRD planes of (111), (200), (220), (311). Bragg's reflection based on the structure of AgNPs which can be indexed consistent with the facets of face-centred spherical crystal nature of silver.

Cytotoxicity of Ps-AgNPs *on TPC1 cells*

Figure 4 shows the cytotoxic efficacy of *Ps*-AgNPs on TPC1 cells was evaluated and carefully interpreted. The results showed that the concentration established inhibition of cell proliferation was observed. The 50% growth inhibition (IC50) was calculated as 48.57 µg/ml of *Ps*-AgNPs. Hence, we selected the dose range of 25, 50 and 75 µg for further experiment.

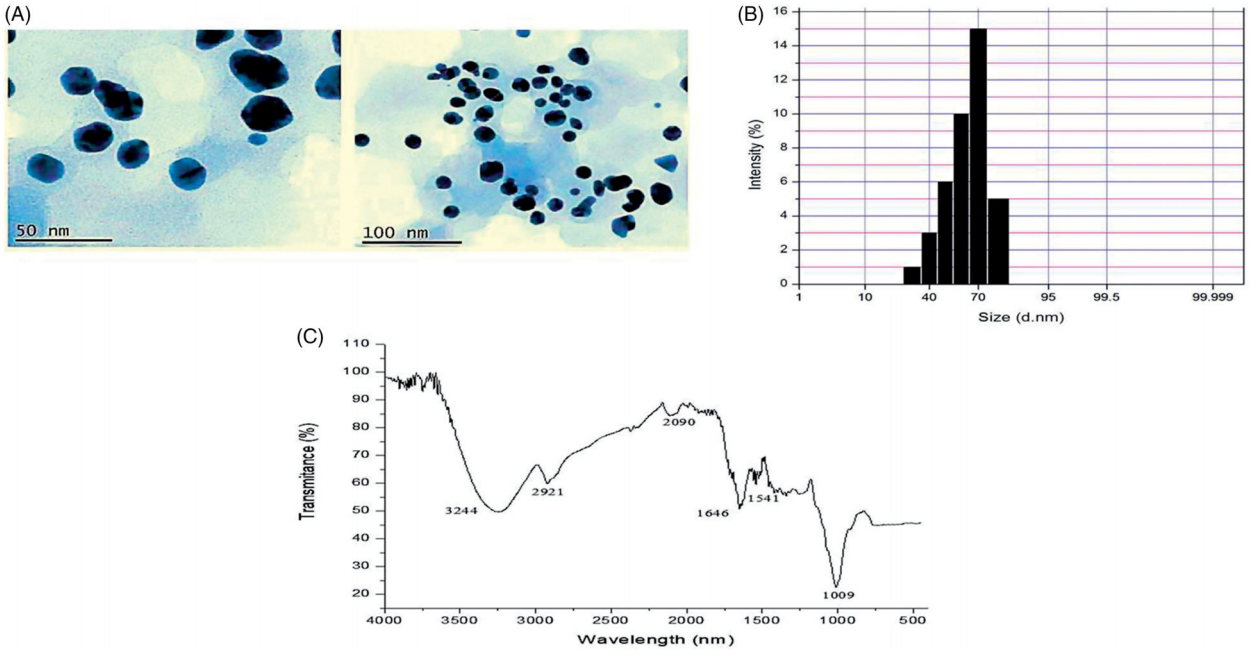


Figure 2. (A) Transmission electron microscopy (TEM) examination images provide the size and morphological characteristics of Ps-AgNPs. (B) Dynamic light scattering measurements show particle size distribution of synthesized Ps-AgNPs. (C) FTIR spectral analyses of various functional groups responsible for the reduction of silver nitrate into silver nanoparticles.

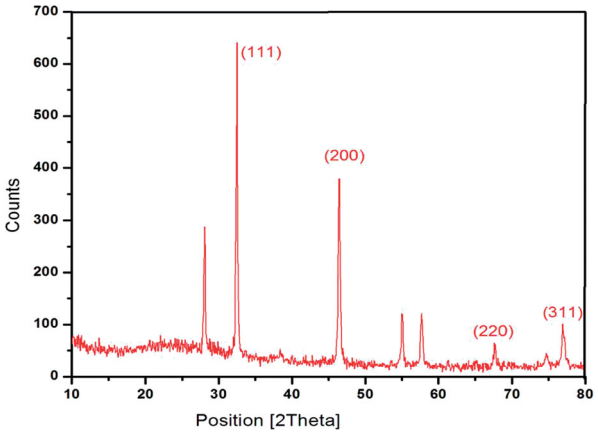


Figure 3. XRD analysis confirmed the crystalline nature of Ps-AgNPs.

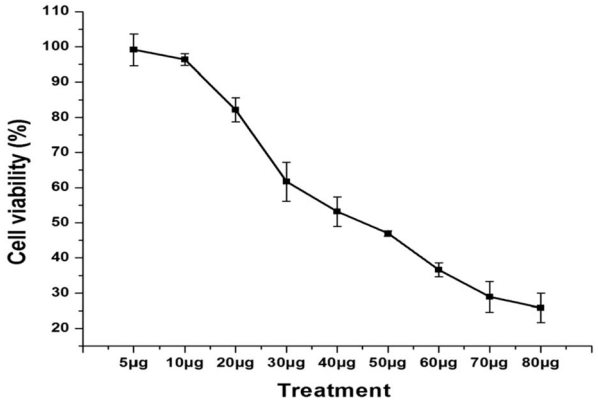


Figure 4. The cytotoxic effect of Ps-AgNPs on TPC1 cells assessed by MTT assay. The statistical analysis was carried out using one way ANOVA. Values are represented as mean ± SD of three experiments. $p < .05$ was significantly different from the control sample.

Effect of Ps-AgNPs on lipid-peroxidation and antioxidant status in TPC1 cells

Figure 5 reveals the levels of lipid peroxidative marker (TBARS) and ranges of antioxidant enzymes (SOD, CAT and GPx) in *Ps*-AgNPs treated TPC1 cells in concentration established way. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs corresponds with the TBARS values of 6.60 ± 0.28, 10.62 ± 0.88 and 14.85 ± 0.20 mM/mg protein, respectively. The untreated control corresponds with the TBARS value of 3.38 ± 0.35 mM/mg protein. It denoted that the level of TBARS was significantly increased in concentration established way in TPC1 cells after treating with *Ps*-AgNPs (25, 50 and 75 µg). However, *Ps*-AgNPs (25, 50 and 75 µg) treatment on TPC1 cells shows decreased levels of SOD, CAT and GPx in concentration established way. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs corresponds with the SOD values of 11.29 ± 0.32, 7.27 ± 0.32 and 4.51 ± 0.49 units/mg protein, respectively. The untreated control showed the SOD level of 16.90 ± 0.35 units/mg protein. The mean CAT values were recorded as 8.43 ± 0.44, 6.82 ± 0.33 and 4.39 ± 0.36 units/mg protein, respectively, and 12.13 ± 0.39 units/mg protein in untreated control. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs correspond with the GPx values of 11.20 ± 1.01, 8.10 ± 0.72 and 4.84 ± 0.58 units/mg protein, respectively, and in untreated control cells, the GPx level was recorded as 15.80 ± 0.77 units/mg protein.

Effect of Ps-AgNPs on intracellular ROS production in TPC1 cells

The ROS generation in TPC1 cells was estimated by exposing the cells with DCFH-DA non-fluorescent probe. Figure 6(A)

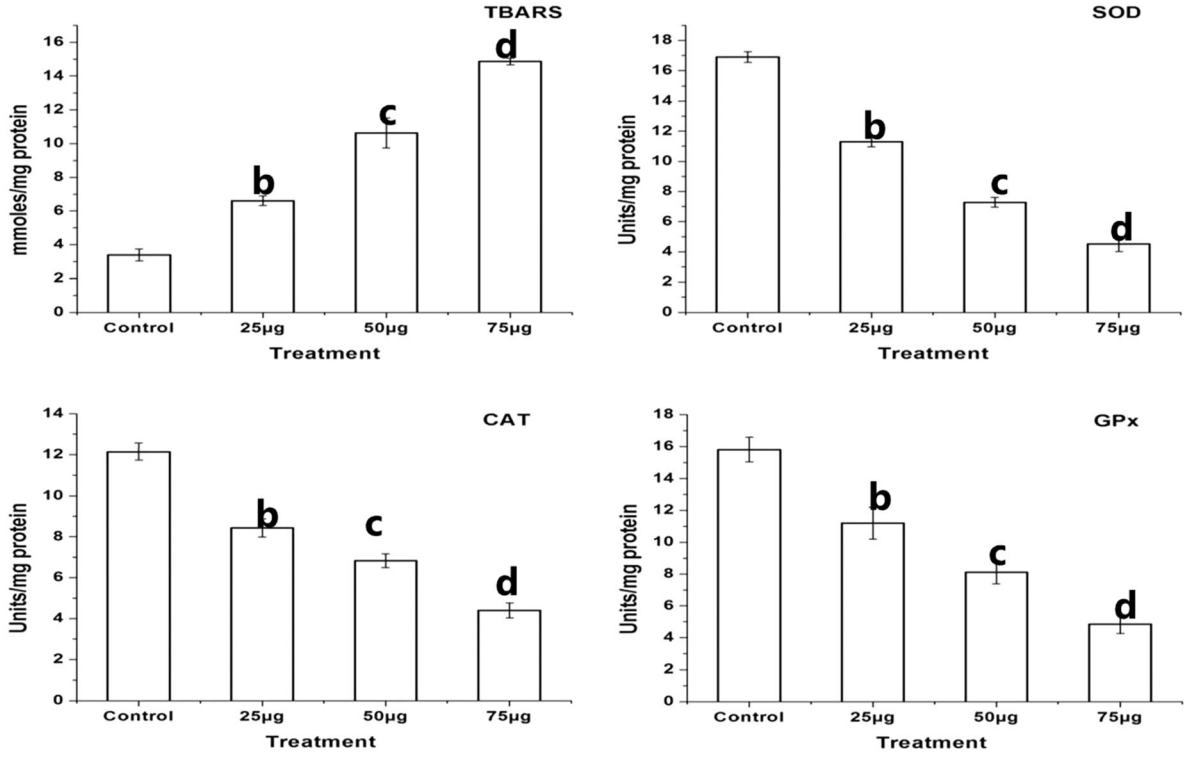


Figure 5. Effect of Ps-AgNPs induced lipid peroxidation and modulate cellular antioxidant status in TPC1 cells. Bars represent the increased TBARS and changes in the activities of SOD, CAT and GPx in Ps-AgNPs (25, 50 and 75 µg) treated TPC1 cells. Values are represented as mean ± SD of three experiments. The b,c,d superscripts indicate significant difference from control ($p < .05$).

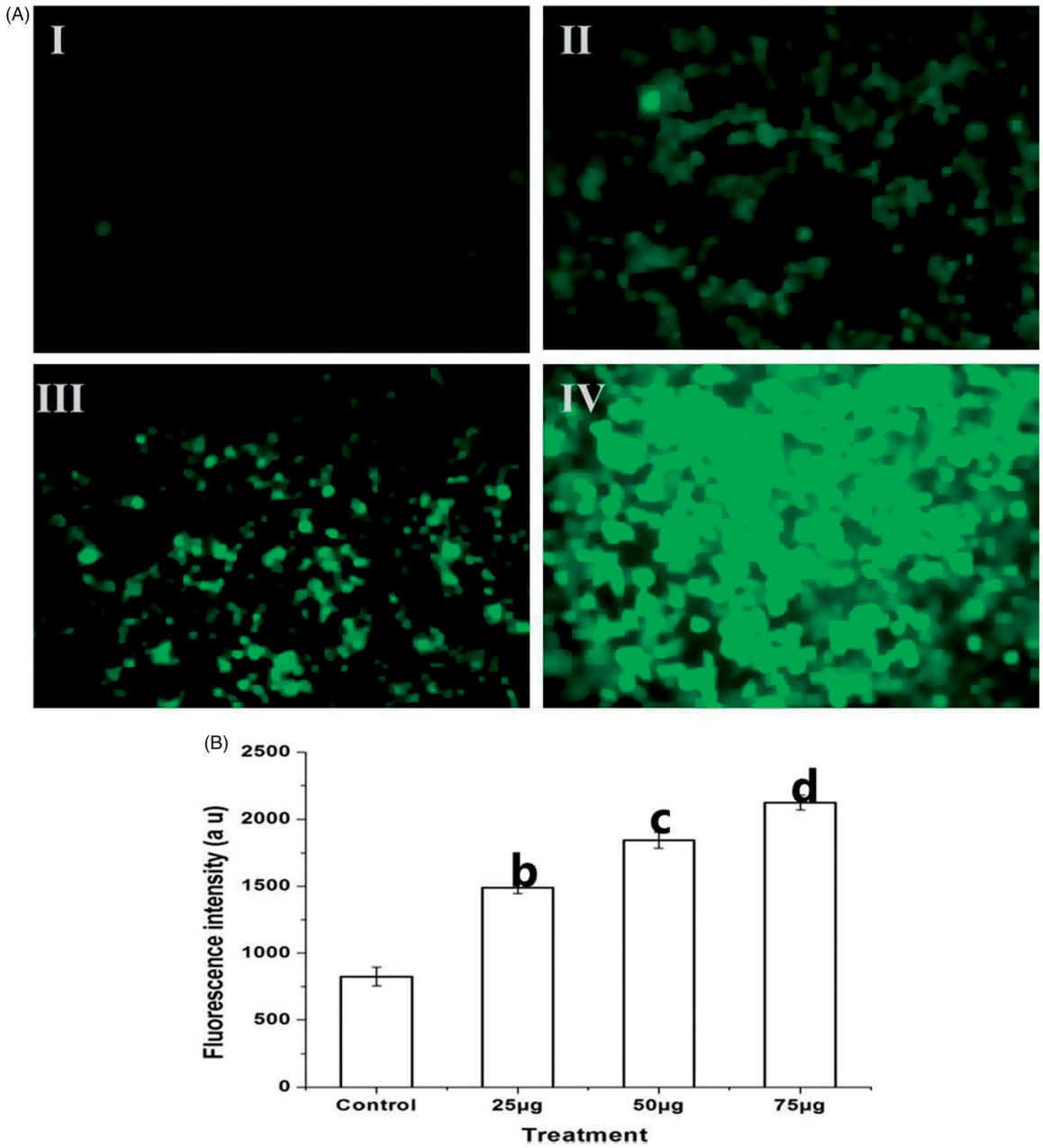


Figure 6. Ps-AgNPs induced intracellular ROS generation in TPC1 cells was evaluated by DCFH-DA staining. (A) Fluorescence images show the control (I) and Ps-AgNPs treated (II–IV) TPC1 cells. The images were acquired by floid cell imaging station. (B) The graphical presentation of intracellular ROS production was detected by spectrofluorometer. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p < .05$).

explores the amounts of ROS production in untreated (I) and cells treated (II–IV) with different concentrations of *Ps*-AgNPs. The observation reveals that the concentration dependent increase of green fluorescence in *Ps*-AgNPs treated cells, denotes the increased intracellular ROS levels which induced by *Ps*-AgNPs in concentration established way. Figure 6(B) illustrates the graphical representation of intracellular ROS production in control and *Ps*-AgNPs treated TPC1.

Effect of Ps-*AgNPs on mitochondrial membrane potential*

By altering MMP, the apoptosis was elicited in *Ps*-AgNPs treated TPC1 cells in a concentration established way, and it was determined by staining the cells with lipophilic cationic dye Rh-123. The increased depth of fluorescence in untreated (I) cells representing polarized mitochondria membrane. Rather, the significantly reduced fluorescence depth was observed in TPC1 cells (II–IV) treated with *Ps*-AgNPs (25, 50 and 75 µg) for 24 h, exhibited significant alteration in mitochondria membrane which leads to cell death (Figure 7(A)). Figure 7(B) shows the graphical representation of the MMP of TPC1.

Effect of Ps-*AgNPs on apoptotic morphological changes in TPC1 cells*

Fluorescence images show the characteristic features of apoptosis in *Ps*-AgNPs treated TPC1 cells stained with

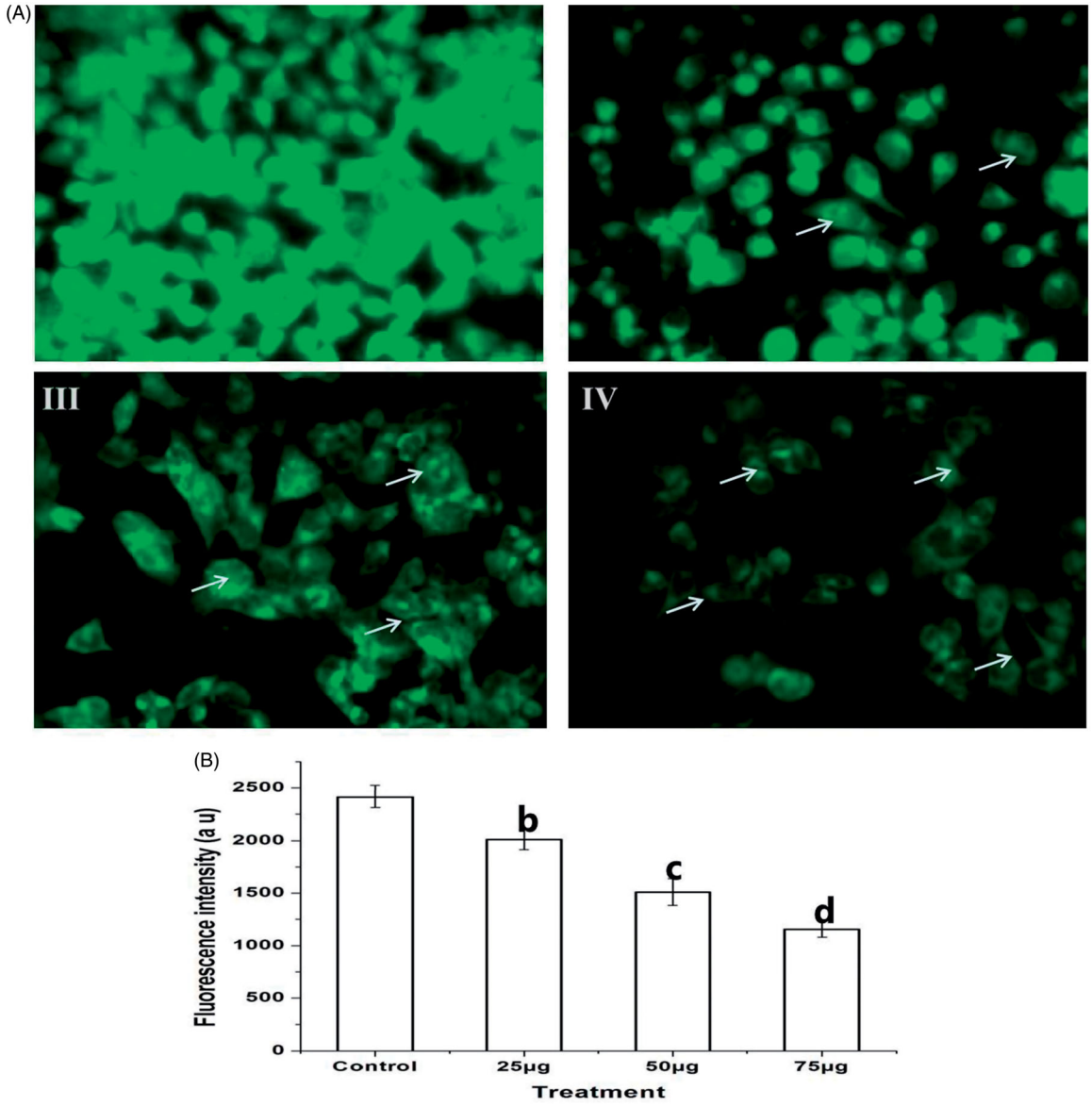


Figure 7. (A) The effect of Ps-AgNPs on the mitochondria membrane potential of TPC1 cells was evaluated by the Rhodamine 123 staining. The images were acquired by floid cell imaging station. The white colour arrow indicates reduced mitochondrial membrane potential with less fluorescence intensity (II–IV). (B) The graphical presentation of mitochondrial membrane potential was detected by spectrofluorometer. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p < .05$).

AO/EtBr (Figure 8(A)). The fluorescence dye AO had only stained the healthy viable cells, whereas the EtBr was selectively entered into condensed nuclei of apoptotic cells. Here, the untreated TPC1 cells had highly green fluorescence nucleus indicating live cells. While the *Ps*-AgNPs treated cells (II–IV) show that orange colour highlights early apoptotic cells and red-stained fragmented nuclei point out late apoptosis cells in concentration established way. Cell condensation in *Ps*-AgNPs treated cells indicates induced apoptotic morphological changes. Figure 8(B) illustrates the graphical representation of *Ps*-AgNPs induced apoptosis in TPC1.

Discussion

Biosynthesis of silver nanoparticle

Microbial source to generate the AgNPs suggests tremendous attention in the nano-biotechnology due to its metabolic activity. It is well known that the production of extracellular enzymes like nitrate reductase by microorganisms is a key source for the reduction of metal ions to nanoparticles [31]. Based on this hypothesize, several works have been conducted to screening genetically engineering special microbial species or silver tolerant bacteria [32,33]. These new insights have challenged these long-held viewpoints that microorganisms are principally accountable for the synthesis of AgNPs in a biologically acceptable way. It was safe to confirm that the biomass including fungi and bacteria catalyzes the AgNPs production only when the reaction is accompanied in the broth-free system and in the absence of light [34].

Here, the bacterial strain *Pseudomonas aeruginosa* was potentially intended for the biosynthesis of AgNPs. Abo-State and Partila [35] evidenced that the *Pseudomonas aeruginosa* was a potential reducer of silver nitrate to AgNPs. Additionally, Kalimuthu et al. [36] reported that AgNPs producing *Bacillus licheniformis* strains showed nitrate reductase

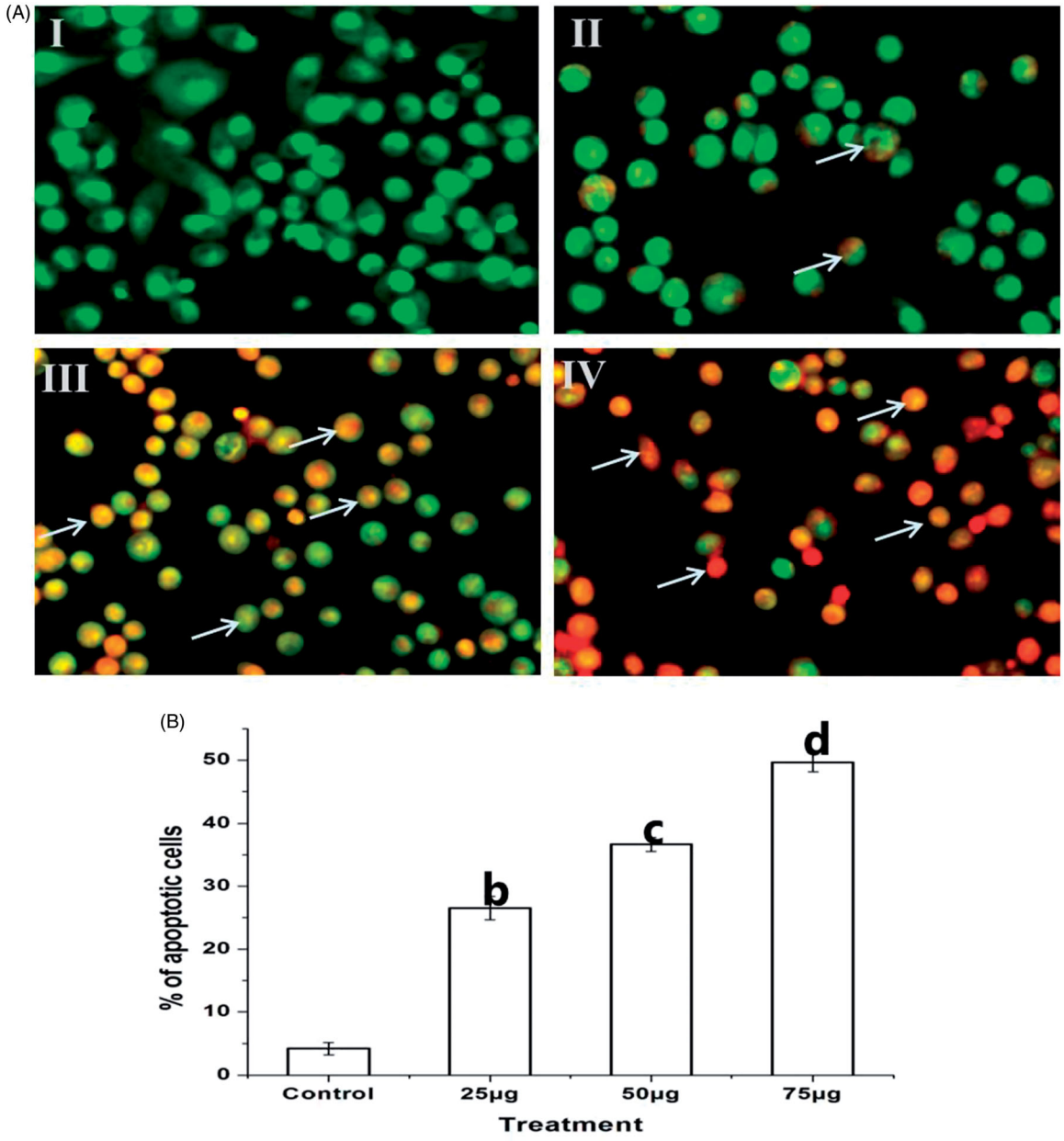


Figure 8. The fluorescence microscopic images illustrate the Ps-AgNPs induced apoptotic morphological changes evaluated by dual staining (AO/EtBr). (A) Untreated TPC1 cells (I) showed green fluorescence indicates viable cells stained with AO. A red colour arrow indicates viable cells and a white colour arrow indicates apoptotic cells in treated cells (II–IV). The images were acquired by floid cell imaging station. (B) The graphical presentation shows the percentage of apoptotic cells. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p<.05$).

activity. According to Thakkar et al. [37], the negatively charged bacterial cell wall interacts with metal ions (positively charged) and then subsequently reduces them into nanoparticles. The conversion of silver ions into AgNPs by microorganisms was described as a defence mechanism to get rid of the toxicity of the metal.

Characterization of Ps-*AgNPs*

Previous reports revealed that the UV spectral peak located at 410 and 450 nm has been witnessed for the presence of AgNPs and might be contributed to sphere-shaped nanoparticles [38]. It was an absolute agreement with UV spectrum of *Ps*-AgNPs obtained in this study. TEM observation of *Ps*-AgNPs was revealed that spherical in form and size was ranged from 20 to 60 nm in diameter it was concurrently agreed with previous reports [7,39]. The existence of the peaks in FT-IR spectrum at 3244 cm^{-1} and 2921 cm^{-1} proportional to a strong acquaintance between O–H and C–H stretching of alcohol, phenol and aromatic compounds [40]. These functional compounds may have an effective function in the biogenic generation of AgNPs. But, the viable mechanism is unclear and needs further investigation.

Cytotoxicity of Ps-*AgNPs on TPC1 cells*

The MTT colorimetric assay was primarily based on the reduction of MTT reagent into formazan by the action of mitochondrial dehydrogenase enzyme which was found in viable cells [41]. The cytotoxic effect of the *Ps*-AgNPs on TPC1 cells

suggests that biosynthesized AgNPs could contribute in the search for a new and alternative anticancer agent. Our findings showed that more than 50% of cell death was achieved at the concentration of 48.57 μg/ml of *Ps*-AgNPs. The cell apoptosis prompted by *Ps*-AgNPs at lower concentrations may be the response of *P. aeruginosa* metabolites accomplished with the *Ps*-AgNPs [42].

The cytotoxicity results of this study were well evidenced that the cytotoxic efficiency of biosynthesized AgNPs counter to liver and colorectal cancer cell line [19]. Metal nanoparticles are multipurpose agents with a variety of biological applications particularly in nanomedicine against multidrug-resistant bacterial pathogens [43] and for various cancerous cells [44]. Hussain et al. [45] investigated the cytotoxic nature of several biosynthesized nanoparticles, including AgNP (15 and 100 nm in size) against liver-derived cancer cells from rats. The 24 h exposure of AgNPs to TPC1 cells, the mitochondrial function and membrane integrity were significantly reduced according to the concentrations used. It has been revealed that exposure of *Ps*-AgNP reduces the ATP production in cells, causes mitochondrial damage and accelerates the ROS production in dose-dependent manner. Several reports suggested that the nanoparticles exposure caused cell cycle arrest in G2/M phase possibly because of damaged DNA repair [46,47].

Conclusions

From the findings, we infer that the *Pseudomonas aeruginosa* is an effective reducer of silver nitrate to the AgNP. This biogenerated AgNPs showed good characteristic features while compared to other methods like phyto-molecule synthesis. It showed effective cytotoxicity on TPC1 cells by inducing intracellular ROS production through increasing lipid peroxidation and decreased antioxidants, reducing the MMP and inducing cell condensation. Through this study, the mechanisms of Ps-AgNPs induced apoptosis in TPC1 cells were well revealed. These results suggested that the microbiological approach is an effective method for nanoparticle fabrication; these biogenic nanoparticles could be used as a novel chemotherapeutic drug for the treatment and/or prevention of thyroid cancer.

Disclosure statement

The authors declare that no conflict of interest.

Funding

This work is supported by The Association Foundation Program of Yunnan Provincial Science and Technology, Department and Kunming Medical University under grant Number [2017FE468(-086)]. The author would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through the Research Group Project No. RGP-300.

References

[1] Gurunathan S, Park JH, Han JW, et al. Comparative assessment of the apoptotic potential of silver nanoparticles synthesized by *Bacillus tequilensis* and *Calocybe indica* in MDA-MB-231 human breast cancer cells: targeting p53 for anticancer therapy. Int J Nanomedicine. 2015;10:4203–4222.

[2] Mukherjee P, Ahmad A, Mandal D. Fungus-mediated synthesis of silver nanoparticles and their immobilization in the mycelial matrix: a novel biological approach to nanoparticle synthesis. Nano Lett. 2001;1(10):515–519.

[3] Conde J, Doria G, Baptista P. Noble metal nanoparticles applications in cancer. J Drug Deliv. 2012;2012:751075.

[4] Ghodake G, Lim SR, Lee DS. Casein hydrolytic peptides mediated green synthesis of antibacterial silver nanoparticles. Colloids Surf B. 2013;108:147–151.

[5] Tanvir S, Oudet F, Pulvin S, et al. Coenzyme based synthesis of silver nanocrystals. Enzyme Microb Technol. 2012;51(4):231–236.

[6] Liu L, Yang J, Xie J, et al. The potent antimicrobial properties of cell penetrating peptide-conjugated silver nanoparticles with excellent selectivity for Gram-positive bacteria over erythrocytes. Nanoscale. 2013;5(9):3834–3840.

[7] Mittal AK, Chisti Y, Banerjee UC. Synthesis of metallic nanoparticles using plant extracts. Biotechnol Adv. 2013;31(2):346–356.

[8] Tuo Y, Liu G, Zhou J, et al. Microbial formation of palladium nanoparticles by *Geobacter sulfurreducens* for chromate reduction. Bioresour Technol. 2013;133:606–611.

[9] Sathyavathi S, Manjula A, Rajendhran J, et al. Extracellular synthesis and characterization of nickel oxide nanoparticles from *Microbacterium* sp. MRS-1 towards bioremediation of nickel electroplating industrial effluent. Bioresour Technol. 2014;165:270–273.

[10] Natarajan K, Selvaraj SV, Murty R. Microbial production of silver Nanoparticles. Dig J Nanomater Biostruct. 2010;5(1):135–140.

[11] Husseiny MI, El-Aziz MA, Badr Y, et al. Biosynthesis of gold nanoparticles using *Pseudomonas aeruginosa*. Spectrochim Acta A . 2007;67(3–4):1003.

[12] Liu L, Shao Z, Ang M, et al. Are microorganisms indispensable in green microbial nanomaterial synthesis. RSC Adv. 2014;4(28):14564.

[13] Zhang X, Yan S, Tyagi RD, et al. Synthesis of nanoparticles by microorganisms and their application in enhancing microbiological reaction rates. Chemosphere. 2011;82(4):489–494.

[14] Pérez-Herrero E, Fernández-Medarde A. Advanced targeted therapies in cancer: drug nanocarriers, the future of chemotherapy. Eur J Pharm Biopharm. 2015;93:52–79.

[15] Hanley C, Layne J, Punnoose A, et al. Preferential killing of cancer cells and activated human T cells using ZnO nanoparticles. Nanotechnology. 2008;19(29):295103.

[16] El-Baghdady K, Elshatoury E, Abdullah MO, et al. Biogenic production of silver nanoparticles by *Enterobacter cloacae* ISM26. Turk J Biol. 2018;42(4):319–321.

[17] Malhotra A, Dolma K, Kaur N, et al. Biosynthesis of gold and silver nanoparticles using a novel marine strain of *Stenotrophomonas*. Bioresour Technol. 2013;31(142):727.

[18] Suresh AK, Pelletier DA, Wang W, et al. Silver nanocrystallites: biofabrication using *Shewanella oneidensis*, and an evaluation of their comparative toxicity on Gram-negative and Gram-positive bacteria. Environ Sci Technol. 2010;44(13):5210–5215.

[19] Buttacavoli M, Albanese NN, Di Cara G, et al. Anticancer activity of biogenerated silver nanoparticles: an integrated proteomic investigation. Oncotarget. 2017;9(11):9685–9705.

[20] Grant CS. Recurrence of papillary thyroid cancer after optimized surgery. Gland Surg. 2015;4(1):52–62.

[21] Davies L, Morris LGT, Haymart M, et al. American association of clinical endocrinologists and American College of Endocrinology Disease State Clinical Review: the increasing incidence of thyroid cancer. Endocr Pract. 2015;21(6):686–696.

[22] Chow S-M, Law SCK, Chan JKC, et al. Papillary microcarcinoma of the thyroid-prognostic significance of lymph node metastasis and multifocality. Cancer. 2003;98(1):31–40.
[23] Gunaseelan S, Balupillai A, Govindasamy K, et al. Linalool prevents oxidative stress activated protein kinases in single UVB-exposed human skin cells. PLOS One. 2017;12(5):e0176699.
[24] Okhawa H, Ohishi N, Yagi K. Assay of lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal Biochem. 1979;95: 351–358.
[25] Kakkar ZYP, Das B, Viswanathan PN. A modified spectrophotometric assay of superoxide dismutase (SOD). Indian J Biochem Biophys. 1984;21:130–132.
[26] Sinha KA. Colorimetric assay of catalase. Anal Biochem. 1972; 47(2):389–394.
[27] Moron MS, De Pierre JW, Mannervik B. Levels of glutathione, glutathione reductase and glutathione S-transferase in rat lung and liver. Biochim Biophys Acta. 1972;82:67–70.
[28] Jesudason E, Masilamoni EG, Charles Jebaraj EJ, et al. Efficacy of DL-alpha lipoic acid against systemic inflammation induced mice: antioxidant defense system. Mol Cell Biochem. 2008;313(1–2): 113–123.
[29] Johnson LV, Walsh ML, Chen LB. Localization of mitochondria in living cells with rhodamine 123. Proc Natl Acad Sci USA. 1980; 77(2):990–994.
[30] Karthikeyan S, Kanimozhi G, Prasad NR, et al. Radiosensitizing effect of ferulic acid on human cervical carcinoma cells *in vitro*. Toxicol In Vitro. 2011;25(7):1366–1375.
[31] Duran N, Marcarto PD, DeSouza GIH, et al. Antibacterial effect of silver nanoparticles produced by fungal process on textile fabrics and their effluent treatment. J Biomed Nanotechnol. 2007;3: 203–208.
[32] Das SK, Khan MR, Guha AK, et al. Silver-nanobiohybrid material: synthesis, characterization and application in water purification. Bioresour Technol. 2012;124:495–499.
[33] Park TJ, Lee SY, Heo NS, et al. In vivo synthesis of diverse metal nanoparticles by recombinant *Escherichia coli*. Angew Chem Int Ed. 2010;49:7019–7024.
[34] Kaler A, Jain S, Banerjee UC. Green and rapid synthesis of anticancerous silver nanoparticles by *Saccharomyces boulardii* and insight into mechanism of nanoparticle synthesis. BioMed Res Int. 2013; 2013:1.
[35] Abo-State MMA, Partila AM. Microbial production of silver nanoparticles by *Pseudomonas aeruginosa* cell free extract. J Ecol Health Environ. 2015; 3(3):91–98.
[36] Kalimuthu K, Suresh Babu R, Venkataraman D, et al. Biosynthesis of silver nanocrystals by *Bacillus licheniformis*. Colloids Surf B. 2008;66:150–153.
[37] Thakkar KN, Mhatre SS, Parikh RY. Biological synthesis of metallic nanoparticles. Nanomedicine. 2010;6(2):257–262.
[38] Zaheer Z, Rafiuddin. Silver nanoparticles to self-assembled films: green synthesis and characterization. Colloids Surf B: Biointerfaces. 2012;90:48–52.
[39] Singh PK, Bhardwaj K, Dubey P, et al. UV-assisted size sampling and anti-bacterial screening of *Lantana camara* leaf extract synthesized silver nanoparticles. RSC Adv. 2015;5(31):24513–24520.
[40] Kumar P, Senthamilselvi S, Lakshmiprabha A, et al. Synthesis of silver nanoparticles from Sargassum tenerrimum and screening phytochemicals for its anti-bacterial activity. Nano Biomed. Eng. 2012;4(1):12–16.
[41] Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods. 1983;65(1–2):55–63.
[42] Farah MA, Ali MA, Chen SM, et al. Silver nanoparticles synthesized from *Adenium obesum* leaf extract induced DNA damage, apoptosis and autophagy via generation of reactive oxygen species. Colloids Surf B: Biointerfaces. 2016;141:158–169.
[43] Behravan M, Hossein Panahi A, Naghizadeh A, et al. Facile green synthesis of silver nanoparticles using *Berberis vulgaris* leaf and root aqueous extract and its antibacterial activity. Int J Biol Macromol. 2019;124:148–154.
[44] Sankar R, Karthik A, Prabu A, et al. *Origanum vulgare* mediated biosynthesis of silver nanoparticles for its antibacterial and anticancer activity. Colloids Surf B Biointerfaces. 2013;1(108):80–84.
[45] Hussain SM, Hess KL, Gearhart JM, et al. In-vitro toxicity of nanoparticles in BRL 3A rat liver cells. Toxicol In Vitro. 2005;19(7): 975–983.
[46] Lee YS, Kim DW, Lee YH, et al. Silver nanoparticles induce apoptosis and G2/M arrest via PKCζ-dependent signaling in A549 lung cells. Arch Toxicol. 2011;85:1529–1540.
[47] AshaRani PV, Low Kah Mun G, Hande MP, et al. Cytotoxicity and genotoxicity of silver nanoparticles in human cells. ACS Nano. 2009;3(2):279–290.

ARTIFICIAL CELLS, NANOMEDICINE, AND BIOTECHNOLOGY
2019, VOL. 48, NO. 1, 800–809
https://doi.org/10.1080/21691401.2019.1687495

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Pseudomonas aeruginosa synthesized silver nanoparticles inhibit cell proliferation and induce ROS mediated apoptosis in thyroid cancer cell line (TPC1)

Jinmei Yang[a], Qiang Wang[b], Chunxiao Wang[c], Ruiyi Yang[a], Mukhtar Ahmed[d], Subramanian Kumaran[e], Periyannan Velu[f] and Bo Li[c]

[a]Physical Examination Center, The First Affiliated Hospital of Kunming Medical University, Kunming, China; [b]Physical Examination Center, The Fourth Affiliated Hospital of Kunming Medical University, Kunming, China; [c]Department of General Surgery, The Fourth Affiliated Hospital of Kunming Medical University, Kunming, China; [d]Department of Zoology, College of Science, King Saud University, Riyadh, Kingdom of Saudi Arabia; [e]Department of Biotechnology, School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology (Deemed to Be University), Chennai, India; [f]Scigen Research and Innovation Pvt. Ltd, Periyar Technology Business Incubator, Thanjavur, India

ABSTRACT

We used cell-free culture filtrate of *Pseudomonas aeruginosa* as a reducing mediator of $AgNO_3$ to silver-nanoparticles (AgNPs) and possibly used as a potential anticancer agent against thyroid cancer cells (TPC1). The bio-generation of AgNPs was firmly established by taking a UV spectrum at 380–500 nm wavelength. The Fourier transform spectrum analysis reveals the association of alcohol, phenol and aromatic functional groups with *P. aeruginosa* synthesized AgNPs (*Ps*-AgNPs). By observing under transmission electron microscopy (TEM), the size and structure of the *Ps*-AgNPs were characterized as the size was 30–70 nm and spherical in shape. The concentration-dependent cytotoxicity of *Ps*-AgNPs on TPC1 cells was observed and IC50 value was calculated. The alteration of oxidative and antioxidant biomarkers in Ps-AgNPs treated cells were observed. The induced apoptosis was determined by staining the *Ps*-AgNPs treated cells with DCFH-DA, Rh-123 dye, Acridine Orange (AO) and ethidium bromide (EtBr). Increased level of intracellular reactive oxygen species (ROS) generation and decreased level of mitochondrial membrane potential was observed in *Ps*-AgNPs treated TPC1 cells. Moreover, the apoptotic morphological changes were explored, which indicates increased apoptosis by inducing cell membrane damage in *Ps*-AgNPs treated cells. This biogenic approach will enable an effective and significant improvement in nano-oncotherapy.

ARTICLE HISTORY
Received 20 May 2019
Revised 7 August 2019
Accepted 10 August 2019

KEYWORDS
Silver nanoparticles; thyroid cancer; TPC1 cells; lipid peroxidation; nano-oncotherapy

Introduction

Nanoparticles synthesis is an important aspect in nanotechnology that has attained growing curious in nano-medicine technology as a result of their potential application in chemotherapies. AgNPs are broadly used in numerous fields like health care, biomedical, food and industrial purposes [1,2]. Nanotechnology tools provide an effective approach against cancers to overcome biological barriers to directly deliver therapeutic drugs [3]. The unique physicochemical properties metallic nanoparticles such as broad optical properties, high surface–volume ratio, ease of synthesis and surface functionalization provide new insights in chemotherapeutics.

Usually, the processes involved in AgNPs synthesis has detailed so far related to the high energy consumption and involvement of hazardous toxic chemicals. The pure biological molecules (like peptide and enzymes) aided AgNPs synthesis has been used successfully around the world [4–6]. This biomolecule approach in nanotechnology has a greater attraction in the field of nanoparticle aided chemotherapies. However, the green reactants approach for AgNPs synthesis based on algal and plant extracts significantly gained great interest due to its biocompatibility, eco-friendliness and benignancy. These appealing features are essential in medical applications. The other advantages of green synthesis include the well-defined and controlled size of the nanoparticles. They are devoid of contaminants and the process is easy to scale-up [7]. Moreover, microorganisms including bacteria, yeast and fungi are being swiftly subjugated as nano-factories in generating numerous nanomaterials [8,9].

The synthesis of nanomaterial by means of the microbial process provides an extensive variety of green and environmentally desirable technique; low-fee production and much less time required [10]. Therefore, the yield effective, harmless, consistent, clean and environmentally friendly methods have required for the preparation of the nanoparticles. The

CONTACT Bo Li lb18669179389@sina.com Department of General Surgery, The Fourth Affiliated Hospital of Kunming Medical University, Kunming 650021, Yunnan, China

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

microorganism for the production of silver nanoparticles (AgNPs) displays great interest in the formation of AgNPs because of its resistant potential to the metals. Synthesis of nanoparticles using cell filtrate could be efficient than the intracellular synthesis [11]. The microbial growth medium contains peptone, yeast extract, dextrose and other important growth factors, some of which may have potential reducing and stabilizing ability as stated in previous reports [12].

For AgNP synthesis, the microbial culture media was aspirated along with remaining microbial biomass after the complete growth of microorganisms. Either with or without the microbial biomass, the silver salt was added into the aspirated growth media. A specific set of enzymes(s) that are secreted by particular microorganisms are consequential elements to define the biosynthesis of nanoparticles, but the main factors for the formation of nanoparticles are not yet clear [13]. Several anticancer drugs were ineffectual to reach their goal due to insufficient concentrations and capably employ the pharmacological consequence without giving harmfulness to healthy tissues and cells [14,15]. Biosynthesized AgNPs, either it was generated by intra- or extracellular mechanisms have been effectively examined as a result of its potent antimicrobial ability against various microbial pathogens [16–18] and cytotoxic nature against various cancer cells [19].

Thyroid cancer is the most common endocrine malignancies among the other cancers. The incidence rate of thyroid cancer has stably increased over the past decade, which considered as one of the causes of death [20]. Papillary thyroid cancer (PTC) is the major histological type which accounts for approximately 80% of human thyroid cancers [21]. The prognosis of PTC has been proven to rely on several well-established clinicopathologic indicators [22]. In this study, we assessed the anticancer potential of *Pseudomonas aeruginosa* generated AgNPs against thyroid cancer cells (TPC1) and then revealed its oxidative stress-induced apoptosis.

Materials and methods

Biosynthesis of AgNPs

For biosynthesis of AgNPs, *Pseudomonas aeruginosa* bacterial strain was periodically cultured and maintained in Nutrient Agar (NA) slants. From 48 h old *P. aeruginosa* culture, one loop full of inoculum was thoroughly suspended in 5 ml of double-distilled water then the suspension was transferred directly to flask containing 45 ml of MGYP (malt extract 3 g, glucose 10 g, yeast extract 3 g and peptone 5 g in 1 l of distilled water and pH 7) broth medium, then the cells were subsequently grown for 48 h in orbital shaking incubator (200 rpm) at 27 °C. For AgNPs synthesis, the cell pellet was harvested by centrifuging the culture broth for 5 min at 6000 rpm and then eroded using 50 mM phosphate buffer (pH 7). The obtained cell pellet was thoroughly suspended in double distilled water then incubated in an orbital shaker at 200 rpm for 24 h. After incubation, the cell-free supernatant was obtained by centrifuging at 7500 rpm for 10 min. Finally, 5 ml of 1 mM $AgNO_3$ solution was transferred to 100 ml flask having 50 ml of collected cell-free culture supernatant, and then the conical flask was effectively cultured at 27 °C in an orbital shaker at 200 rpm for 20 min [23].

Evaluation of maximum AgNPs synthesis

In order to evaluate the maximum synthesis of nanoparticle at different incubation period (1 h, 24 h, 7 days and 15 days), totally four reactions (cell-free culture supernatant and $AgNO_3$) were prepared as described above. The absorption spectral readings of reactants were recorded by ultra violet–visible (UV–vis) spectrophotometer with 0.72 nm resolution (Shimadzu, Kyoto, Japan, UV Pharma spec 1700) from 300 to 800 nm.

Characterization of biosynthesized Ps-AgNPs

The size and structural features of the biosynthesized *P. aeruginosa* synthesized AgNPs (*Ps*-AgNPs) were characterized by transmission electron microscopy (TEM, JEOL-JEM-2010, Tokyo, Japan). The crystalline nature of the *Ps*-AgNPs was examined by X-ray diffraction analyzer (D/Max 2005, Rigaku, Tokyo, Japan). In order to explore the surface functional groups associated with synthesized AgNPs, the *Ps*-AgNPs subjected to Fourier transform infrared spectrophotometer analysis (FT-IR; Shimadzu 8201PC, Kyoto, Japan).

Cancer cell culture

The thyroid cancer cells (TPC1) were inoculated in 10% foetal bovine serum and 1% antibiotics (penicillin, streptomycin, and amphotericin) supplied Dulbecco's modified Eagle's medium (DMEM), incubated at 37 °C in a moistened atmosphere provided with 5% CO_2 and 95% air incubation. *Ps*-AgNPs were completely dissolved in 1% DMSO prior treatment with cell.

Cytotoxicity of Ps-AgNPs against TPC1 cells

The cytotoxicity of *Ps*-AgNPs was evaluated by employing 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay based on the reduction of MTT by the action of mitochondrial reductase enzyme which presents in viable cells. The amount of reduction of MTT to formazan in *Ps*-AgNPs treated and untreated cells were directly proportional to the number of viable cells, it was assessed by taking absorbance values by dissolving the formazan with DMSO. In brief, 1×10^5 cells were plated into 96-well plates and incubated for 24 h at 37 °C in a humidified condition. After complete incubation, the old medium was aspirated without disturbing cells then the cells were treated with different concentrations (5, 10, 20, 30, 40, 50, 60, 70 and 80 μg/ml) of *Ps*-AgNPs dissolved DMEM, the plate was kept in same incubation condition. After incubation, 100 μl of MTT reagent (5 mg/ml in PBS) was subsequently added into each well further, the plates were incubated for 24 h at 37 °C. The resulting formazan was dissolved by adding 100 μl of DMSO and the absorbance of the reactant solution was recorded at 595 nm wavelength using a multiwell plate reader (Tecan

Multimode Reader, Grödig, Austria). The concentrations of the test sample which showed 50% of cell death were then calculated [23].

Biochemical assays

The TPC1 cells were harvested after treatment of biosynthesized *Ps*-AgNPs (25, 50 and 75 μg) and subsequently subjected to biochemical assessments. The levels of lipid peroxidation in *Ps*-AgNPs treated cells were estimated by measuring LPO (lipid peroxide) byproduct of TBARS (thiobarbituric acid) reactive substance by the method described by Okhawa et al. [24], and the activities of enzymatic antioxidants such as superoxide dismutase (SOD) [25], catalase (CAT) [26], the intracellular enzyme of reduced glutathione (GPx) level [27] were measured. The untreated well was employed as a control for all assays, respectively.

Measurement of intracellular ROS generation

Intracellular reactive oxygen species (ROS) production in *Ps*-AgNPs treated and untreated TPC1 cells were estimated by non-fluorescent probe DCFH-DA that can easily enter into the intracellular cell matrix, there it was oxidized into fluorescent dichlorofluorescein (DCF) by the action of produced ROS. Accordingly, the fluorescence strength was relatively proportional to the level of ROS production [28]. The TPC1 cells were inoculated (1×10^6 cells/well) into six-well plate, treated with different concentrations (25, 50 and 75 μg) of *Ps*-AgNPs and placed in CO_2 (5%) incubator for 24 h. After incubation, cells were exposed to 100 μl of DCFH-DA for 10 min at 37 °C. Fluorescence depth was estimated with excitation and emission filters set at 485 and 530 nm, respectively (Shimadzu RF-5301 PC spectrofluorometer, Kyoto, Japan). The results showed an increased percentage of fluorescence depth.

Determination of mitochondrial membrane potential

The mitochondrial membrane potential (MMP) of treated and untreated TPC1 cells was estimated by staining the cells with Rhodamine-123 (Rh-123), lipophilic cationic dye [29]. The cells (3×10^4 cells/well) were cultured in six wells plate and treated with various concentrations of *Ps*-AgNPs (25, 50 and 75 μg) for 24 h. Then, Rh-123 fluorescent dye was added to treated and untreated cells and left for 30 min. The MMP was qualitatively examined under a FLoid cell imaging station (Invitrogen, Carlsbad, CA, USA). Subsequently, the cells were harvested by trypsinization of cells and the fluorescence strength was estimated at 485/530 nm wavelength under spectrofluorometer (Shimadzu, Columbia, MD, USA). The results were compared with positive control which maintained without treatment.

Determination of apoptotic morphological changes

Staining the treated and untreated TPC1 cells with Acridine Orange (AO) and ethidium bromide (EtBr) dye was used to determine the apoptotic cells morphological changes by *Ps*-AgNPs [30]. The cells were cultivated in six-well plate (3×10^4 cells/well) and treated with *Ps*-AgNPs at various dosage ranges (25, 50 and 75 μg) for 24 h. The cells were fixed in methanol:glacial acetic acid (3:1) for 30 min at 4 °C. The cells were stained with AO/EtBr (1:1 ratio) and left for 30 min after washing cells with PBS. The stained cells were consequently washed again with PBS and observed under a FLoid cell imaging station (Invitrogen, Carlsbad, CA, USA). The apoptosis was assessed by counting the cells showing apoptosis features in the total number of cells observed in the microscopic field.

Statistical analyses

All experiments were conducted in three independent replicated and the results were stated as the mean ± standard deviation (mean ± SD) analyzing by one-way analysis of variance (ANOVA) between the obtained values. Values of $p < .05$ indicated the significant differences in data.

Results

Biosynthesis of AgNPs

The biogenic generation of AgNPs by the cell-free extract of *P. aeruginosa* was initially confirmed by visual observation as shown in Figure 1(A). It was noticed that the no colour changes in $AgNO_3$ solution alone, which indicates that no convention of AgNPs. The colour change from pale yellow to brown was noticed in a flask containing the culture filtrate and $AgNO_3$, which denoted that the reduction of $AgNO_3$ to AgNPs. The colour change to dark brown was observed in extend incubation period up to 6 h.

Further, the *Ps*-AgNPs synthesis was investigated by taking spectral readings by UV–vis spectrophotometer at 380–500 nm wavelength. Figure 1(B) shows the UV–visible readings of AgNPs synthesized at various time intervals (1 h, 24 h, 7 days and 15 days) after added silver nitrate to cell filtrate resulted in the increasing spectrum of absorption at around 420 nm. The absorption spectrum of *Ps*-AgNPs displayed a greater wide peak at 420 nm and allocated to surface plasmon resonance of the *Ps*-AgNPs. Moreover, the increased spectrum at 420 nm was observed an extended incubation period.

Characterization of Ps-AgNPs

TEM

A TEM examination provides a better understanding of the size and structural features of the *Ps*-AgNPs. The TEM images at various magnifications ranges are depicted in Figure 2(A), which shows the shape and size of the bio-generated *Ps*-AgNPs. It revealed that the *Ps*-AgNPs sphere-shaped and the size were ranged from 30 to 75 nm. As shown in Figure 2(B), the average size of synthesized *Ps*-AgNPs was 70 mm and has 15.1% of occurrence.

Fourier transform infra-red (FTIR) spectroscopy analyses

In order to analyze the numerous functional groups accountable for the reduction of silver nitrate to AgNPs, the FTIR spectrum patterns were taken and subsequently interpreted. As shown in Figure 2(C), the spectrum depicts the number of peaks that were located at 3244, 2921, 2090, 1646, 1541 and 1009 cm^{-1}. The band arising at 3244 cm^{-1} in the spectrum resembles O–H stretching representing the occurrence of alcohol and phenol groups. A band at 2921 cm^{-1} region was assigned for C–H stretching of an aromatic compound. The peak at 2090 cm^{-1} region was arising from N=C=S stretching of isothiocyanate. A peak raised at 1646 cm^{-1} resembles C=C stretching of the conjugated alkene. A band at 1541 cm^{-1} assigned for strong N–O stretching of the nitro compound and the band at 1009 cm^{-1} strong C–F stretching of fluoro compound.

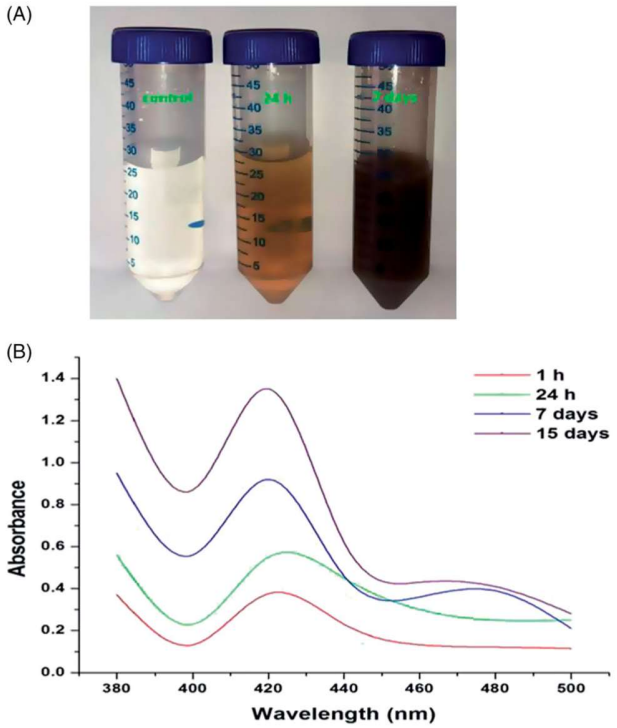


Figure 1. (A) Visual observations of colour change from fine yellow to reddish brown; indicates the reduction of $AgNO_3$ into silver nanoparticles. (B) UV–visible spectrum of biosynthesized silver nanoparticle at different time intervals (1 h, 24 h, 7 days and 15 days) at 380–500 nm. The increased spectrum at 420 nm explores synthesize of silver nanoparticle.

X-ray diffraction analyses

Figure 3 shows the XRD analysis established the crystal-like nature of *Ps*-AgNPs, that depicts peaks at 2θ values of four distinct diffraction peaks at 33.28°, 46.38°, 67.54° and 77.64° corresponding to XRD planes of (111), (200), (220), (311). Bragg's reflection based on the structure of AgNPs which can be indexed consistent with the facets of face-centred spherical crystal nature of silver.

Cytotoxicity of Ps-*AgNPs on TPC1 cells*

Figure 4 shows the cytotoxic efficacy of *Ps*-AgNPs on TPC1 cells was evaluated and carefully interpreted. The results showed that the concentration established inhibition of cell proliferation was observed. The 50% growth inhibition (IC50) was calculated as 48.57 µg/ml of *Ps*-AgNPs. Hence, we selected the dose range of 25, 50 and 75 µg for further experiment.

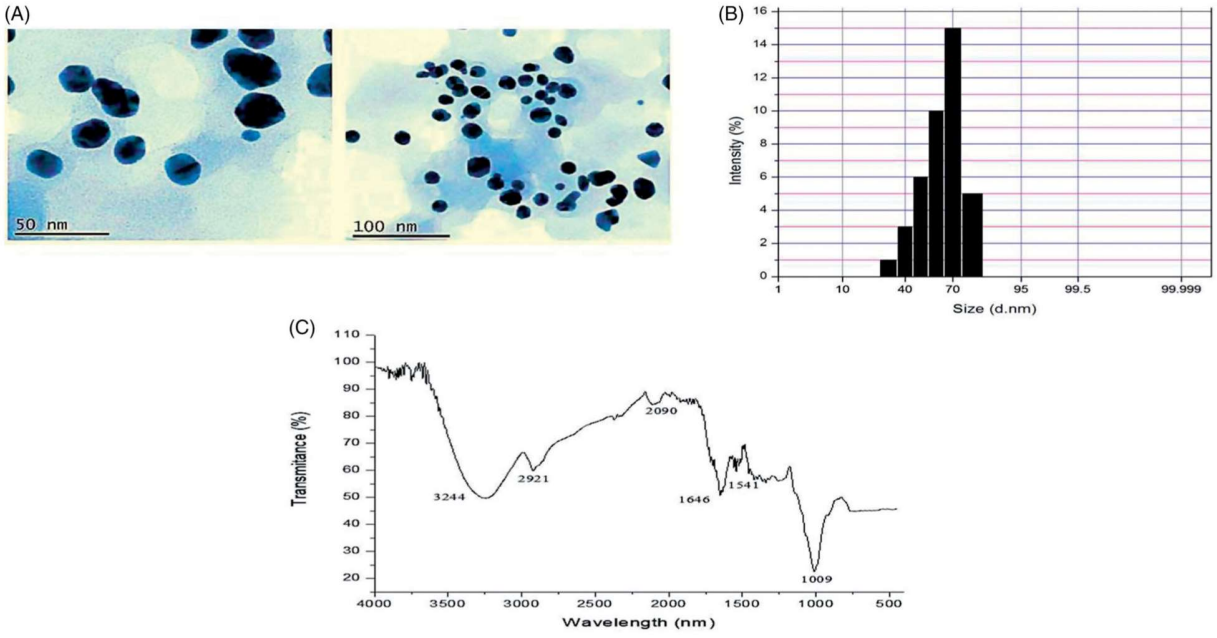


Figure 2. (A) Transmission electron microscopy (TEM) examination images provide the size and morphological characteristics of Ps-AgNPs. (B) Dynamic light scattering measurements show particle size distribution of synthesized Ps-AgNPs. (C) FTIR spectral analyses of various functional groups responsible for the reduction of silver nitrate into silver nanoparticles.

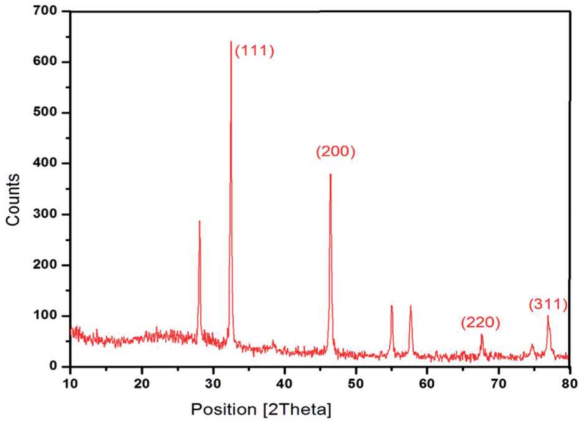


Figure 3. XRD analysis confirmed the crystalline nature of Ps-AgNPs.

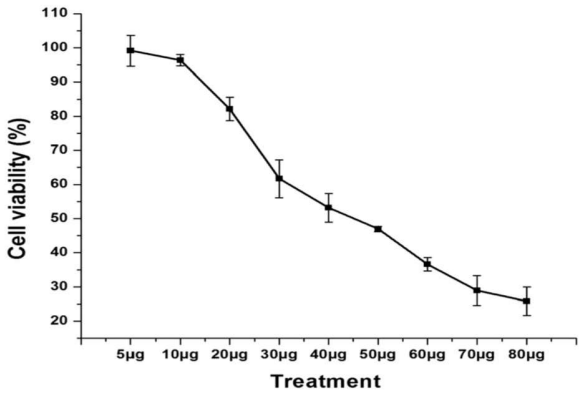


Figure 4. The cytotoxic effect of Ps-AgNPs on TPC1 cells assessed by MTT assay. The statistical analysis was carried out using one way ANOVA. Values are represented as mean ± SD of three experiments. $p<.05$ was significantly different from the control sample.

Effect of Ps-AgNPs on lipid-peroxidation and antioxidant status in TPC1 cells

Figure 5 reveals the levels of lipid peroxidative marker (TBARS) and ranges of antioxidant enzymes (SOD, CAT and GPx) in *Ps*-AgNPs treated TPC1 cells in concentration established way. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs corresponds with the TBARS values of 6.60 ± 0.28, 10.62 ± 0.88 and 14.85 ± 0.20 mM/mg protein, respectively. The untreated control corresponds with the TBARS value of 3.38 ± 0.35 mM/mg protein. It denoted that the level of TBARS was significantly increased in concentration established way in TPC1 cells after treating with *Ps*-AgNPs (25, 50 and 75 µg). However, *Ps*-AgNPs (25, 50 and 75 µg) treatment on TPC1 cells shows decreased levels of SOD, CAT and GPx in concentration established way. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs corresponds with the SOD values of 11.29 ± 0.32, 7.27 ± 0.32 and 4.51 ± 0.49 units/mg protein, respectively. The untreated control showed the SOD level of 16.90 ± 0.35 units/mg protein. The mean CAT values were recorded as 8.43 ± 0.44, 6.82 ± 0.33 and 4.39 ± 0.36 units/mg protein, respectively, and 12.13 ± 0.39 units/mg protein in untreated control. The treatment with 25, 50 and 75 µg of *Ps*-AgNPs correspond with the GPx values of 11.20 ± 1.01, 8.10 ± 0.72 and 4.84 ± 0.58 units/mg protein, respectively, and in untreated control cells, the GPx level was recorded as 15.80 ± 0.77 units/mg protein.

Effect of Ps-AgNPs on intracellular ROS production in TPC1 cells

The ROS generation in TPC1 cells was estimated by exposing the cells with DCFH-DA non-fluorescent probe. Figure 6(A)

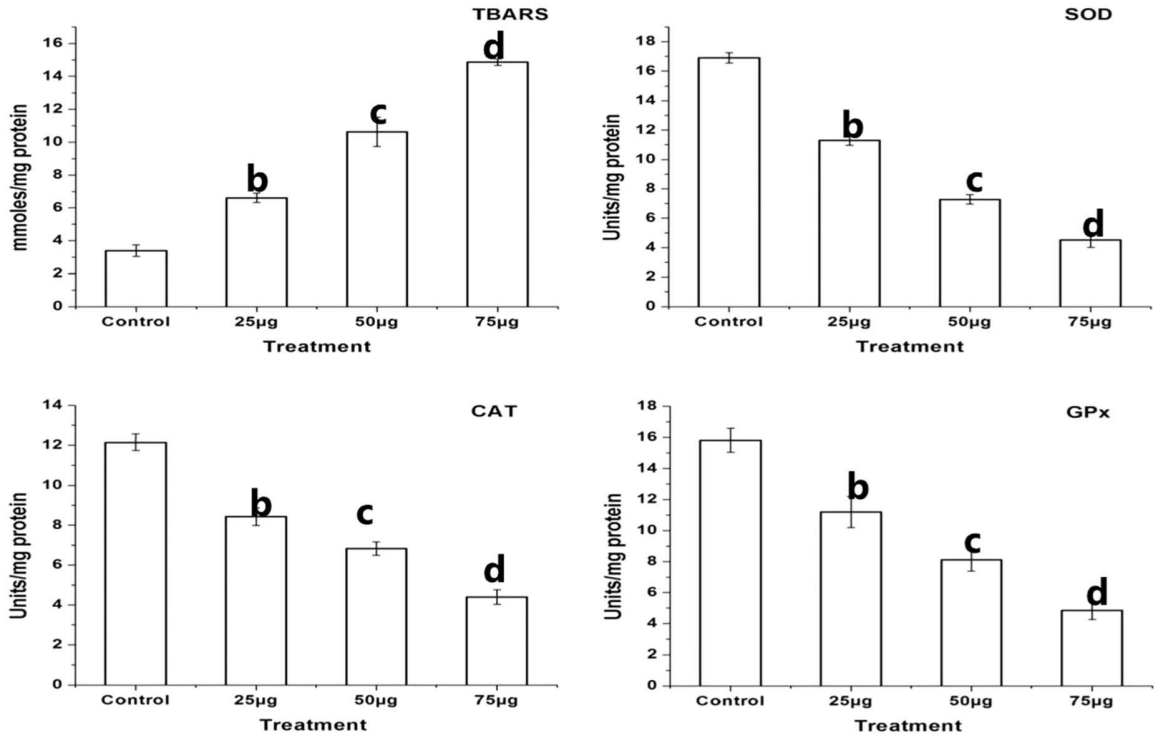


Figure 5. Effect of Ps-AgNPs induced lipid peroxidation and modulate cellular antioxidant status in TPC1 cells. Bars represent the increased TBARS and changes in the activities of SOD, CAT and GPx in Ps-AgNPs (25, 50 and 75 µg) treated TPC1 cells. Values are represented as mean ± SD of three experiments. The b,c,d superscripts indicate significant difference from control ($p < .05$).

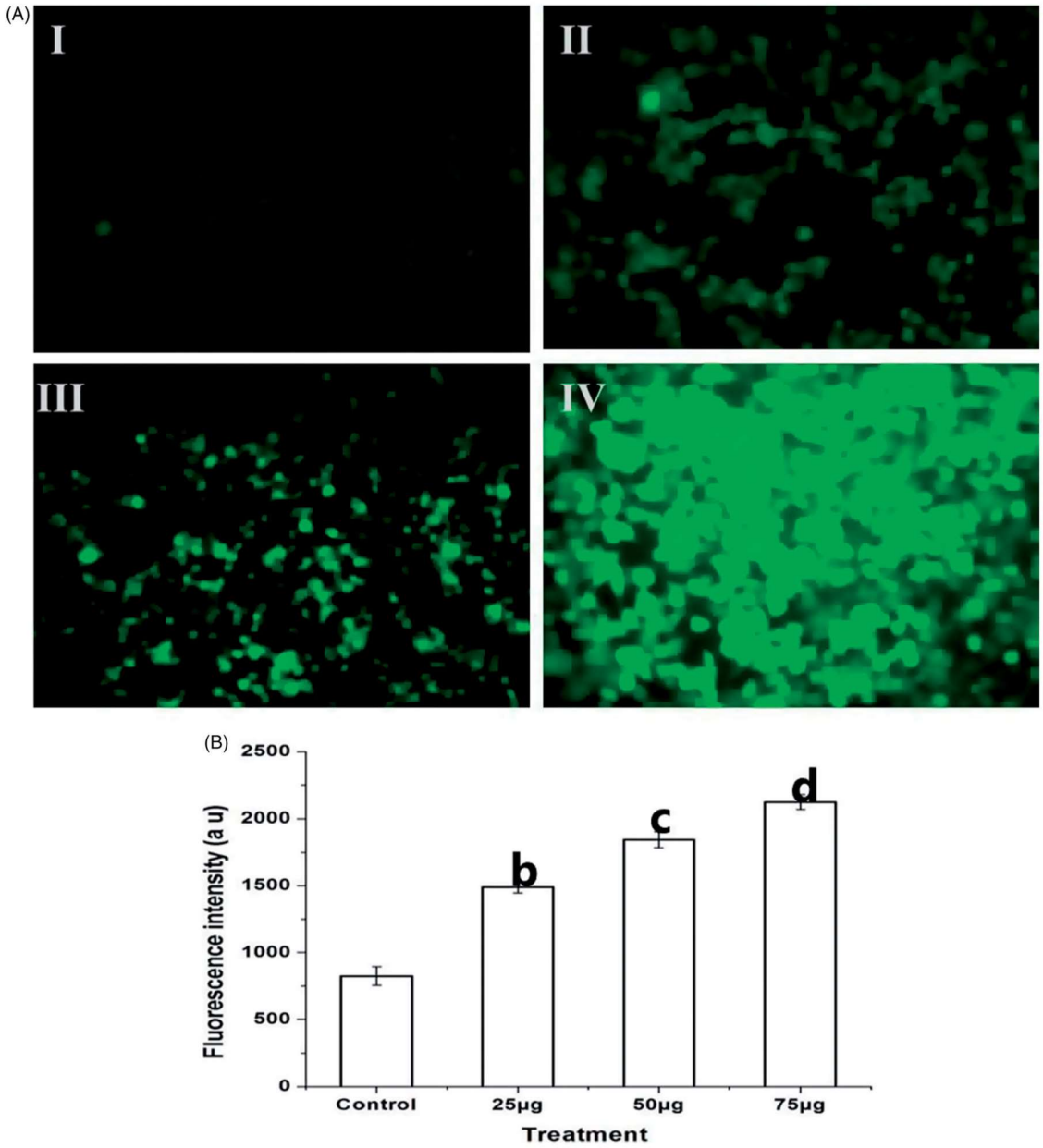


Figure 6. Ps-AgNPs induced intracellular ROS generation in TPC1 cells was evaluated by DCFH-DA staining. (A) Fluorescence images show the control (I) and Ps-AgNPs treated (II–IV) TPC1 cells. The images were acquired by floid cell imaging station. (B) The graphical presentation of intracellular ROS production was detected by spectrofluorometer. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p < .05$).

explores the amounts of ROS production in untreated (I) and cells treated (II–IV) with different concentrations of *Ps*-AgNPs. The observation reveals that the concentration dependent increase of green fluorescence in *Ps*-AgNPs treated cells, denotes the increased intracellular ROS levels which induced by *Ps*-AgNPs in concentration established way. Figure 6(B) illustrates the graphical representation of intracellular ROS production in control and *Ps*-AgNPs treated TPC1.

Effect of Ps-*AgNPs on mitochondrial membrane potential*

By altering MMP, the apoptosis was elicited in *Ps*-AgNPs treated TPC1 cells in a concentration established way, and it was determined by staining the cells with lipophilic cationic dye Rh-123. The increased depth of fluorescence in untreated (I) cells representing polarized mitochondria membrane. Rather, the significantly reduced fluorescence depth was observed in TPC1 cells (II–IV) treated with *Ps*-AgNPs (25, 50 and 75 µg) for 24 h, exhibited significant alteration in mitochondria membrane which leads to cell death (Figure 7(A)). Figure 7(B) shows the graphical representation of the MMP of TPC1.

Effect of Ps-*AgNPs on apoptotic morphological changes in TPC1 cells*

Fluorescence images show the characteristic features of apoptosis in *Ps*-AgNPs treated TPC1 cells stained with

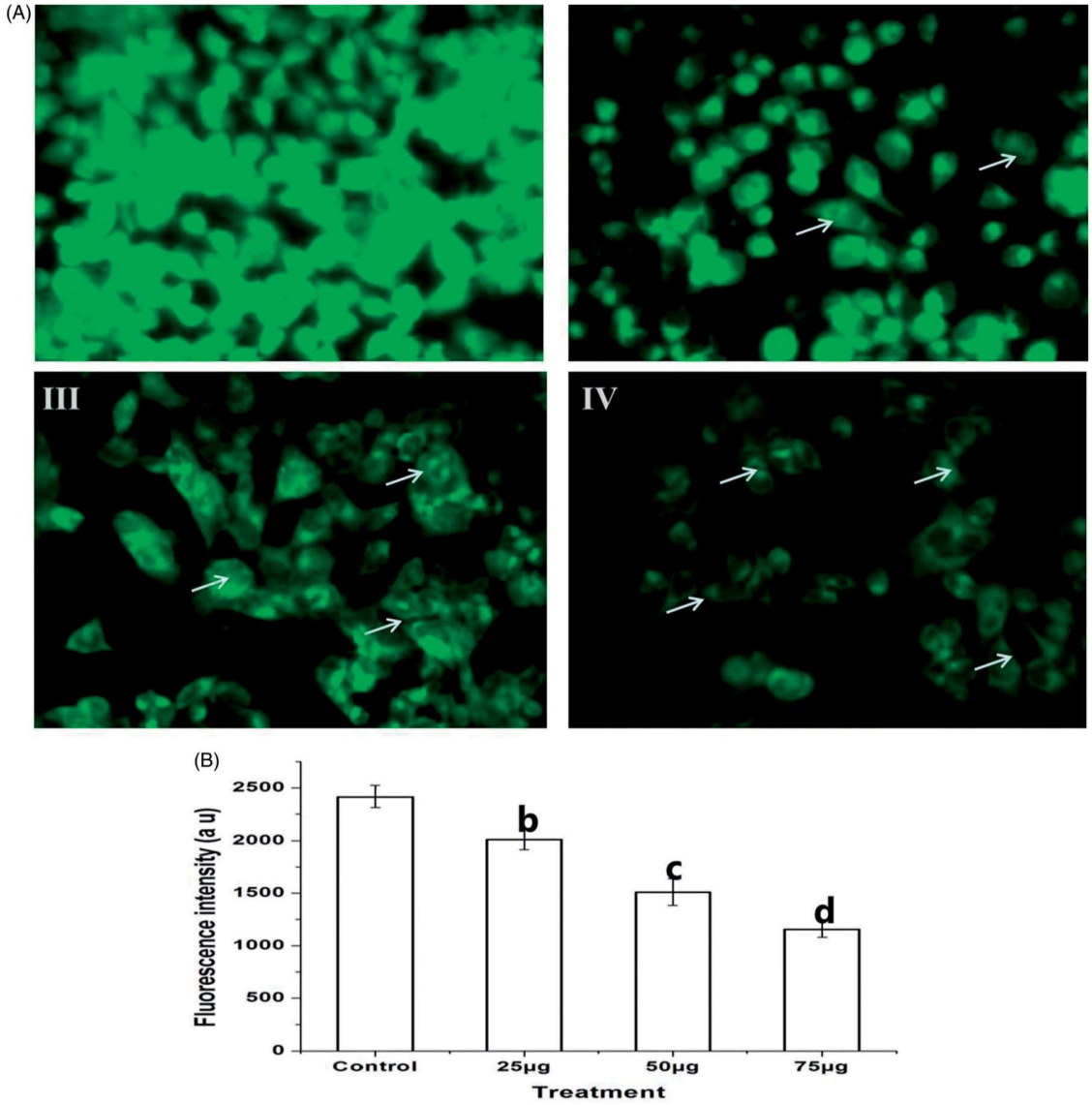


Figure 7. (A) The effect of Ps-AgNPs on the mitochondria membrane potential of TPC1 cells was evaluated by the Rhodamine 123 staining. The images were acquired by floid cell imaging station. The white colour arrow indicates reduced mitochondrial membrane potential with less fluorescence intensity (II–IV). (B) The graphical presentation of mitochondrial membrane potential was detected by spectrofluorometer. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p<.05$).

AO/EtBr (Figure 8(A)). The fluorescence dye AO had only stained the healthy viable cells, whereas the EtBr was selectively entered into condensed nuclei of apoptotic cells. Here, the untreated TPC1 cells had highly green fluorescence nucleus indicating live cells. While the *Ps*-AgNPs treated cells (II–IV) show that orange colour highlights early apoptotic cells and red-stained fragmented nuclei point out late apoptosis cells in concentration established way. Cell condensation in *Ps*-AgNPs treated cells indicates induced apoptotic morphological changes. Figure 8(B) illustrates the graphical representation of *Ps*-AgNPs induced apoptosis in TPC1.

Discussion

Biosynthesis of silver nanoparticle

Microbial source to generate the AgNPs suggests tremendous attention in the nano-biotechnology due to its metabolic activity. It is well known that the production of extracellular enzymes like nitrate reductase by microorganisms is a key source for the reduction of metal ions to nanoparticles [31]. Based on this hypothesize, several works have been conducted to screening genetically engineering special microbial species or silver tolerant bacteria [32,33]. These new insights have challenged these long-held viewpoints that microorganisms are principally accountable for the synthesis of AgNPs in a biologically acceptable way. It was safe to confirm that the biomass including fungi and bacteria catalyzes the AgNPs production only when the reaction is accompanied in the broth-free system and in the absence of light [34].

Here, the bacterial strain *Pseudomonas aeruginosa* was potentially intended for the biosynthesis of AgNPs. Abo-State and Partila [35] evidenced that the *Pseudomonas aeruginosa* was a potential reducer of silver nitrate to AgNPs. Additionally, Kalimuthu et al. [36] reported that AgNPs producing *Bacillus licheniformis* strains showed nitrate reductase

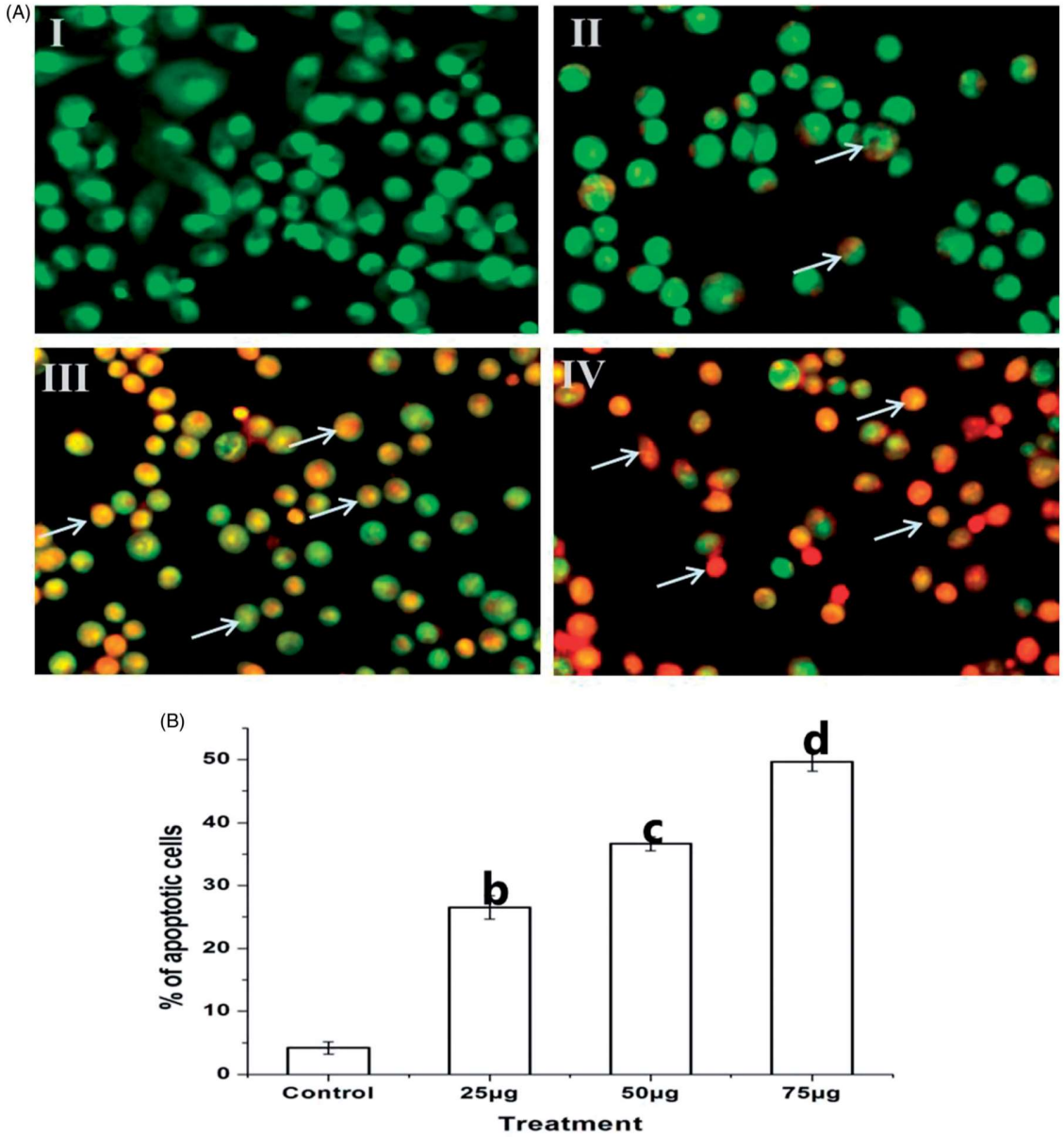


Figure 8. The fluorescence microscopic images illustrate the Ps-AgNPs induced apoptotic morphological changes evaluated by dual staining (AO/EtBr). (A) Untreated TPC1 cells (I) showed green fluorescence indicates viable cells stained with AO. A red colour arrow indicates viable cells and a white colour arrow indicates apoptotic cells in treated cells (II–IV). The images were acquired by floid cell imaging station. (B) The graphical presentation shows the percentage of apoptotic cells. All experiments were performed in triplicate and all the values were expressed as mean ± standard deviation of three replicates. The b,c,d superscripts indicate significant difference from control ($p<.05$).

activity. According to Thakkar et al. [37], the negatively charged bacterial cell wall interacts with metal ions (positively charged) and then subsequently reduces them into nanoparticles. The conversion of silver ions into AgNPs by microorganisms was described as a defence mechanism to get rid of the toxicity of the metal.

Characterization of Ps-*AgNPs*

Previous reports revealed that the UV spectral peak located at 410 and 450 nm has been witnessed for the presence of AgNPs and might be contributed to sphere-shaped nanoparticles [38]. It was an absolute agreement with UV spectrum of *Ps*-AgNPs obtained in this study. TEM observation of *Ps*-AgNPs was revealed that spherical in form and size was ranged from 20 to 60 nm in diameter it was concurrently agreed with previous reports [7,39]. The existence of the peaks in FT-IR spectrum at 3244 cm^{-1} and 2921 cm^{-1} proportional to a strong acquaintance between O–H and C–H stretching of alcohol, phenol and aromatic compounds [40]. These functional compounds may have an effective function in the biogenic generation of AgNPs. But, the viable mechanism is unclear and needs further investigation.

Cytotoxicity of Ps-*AgNPs on TPC1 cells*

The MTT colorimetric assay was primarily based on the reduction of MTT reagent into formazan by the action of mitochondrial dehydrogenase enzyme which was found in viable cells [41]. The cytotoxic effect of the *Ps*-AgNPs on TPC1 cells

suggests that biosynthesized AgNPs could contribute in the search for a new and alternative anticancer agent. Our findings showed that more than 50% of cell death was achieved at the concentration of 48.57 μg/ml of *Ps*-AgNPs. The cell apoptosis prompted by *Ps*-AgNPs at lower concentrations may be the response of *P. aeruginosa* metabolites accomplished with the *Ps*-AgNPs [42].

The cytotoxicity results of this study were well evidenced that the cytotoxic efficiency of biosynthesized AgNPs counter to liver and colorectal cancer cell line [19]. Metal nanoparticles are multipurpose agents with a variety of biological applications particularly in nanomedicine against multidrug-resistant bacterial pathogens [43] and for various cancerous cells [44]. Hussain et al. [45] investigated the cytotoxic nature of several biosynthesized nanoparticles, including AgNP (15 and 100 nm in size) against liver-derived cancer cells from rats. The 24 h exposure of AgNPs to TPC1 cells, the mitochondrial function and membrane integrity were significantly reduced according to the concentrations used. It has been revealed that exposure of *Ps*-AgNP reduces the ATP production in cells, causes mitochondrial damage and accelerates the ROS production in dose-dependent manner. Several reports suggested that the nanoparticles exposure caused cell cycle arrest in G2/M phase possibly because of damaged DNA repair [46,47].

Conclusions

From the findings, we infer that the *Pseudomonas aeruginosa* is an effective reducer of silver nitrate to the AgNP. This biogenerated AgNPs showed good characteristic features while compared to other methods like phyto-molecule synthesis. It showed effective cytotoxicity on TPC1 cells by inducing intracellular ROS production through increasing lipid peroxidation and decreased antioxidants, reducing the MMP and inducing cell condensation. Through this study, the mechanisms of Ps-AgNPs induced apoptosis in TPC1 cells were well revealed. These results suggested that the microbiological approach is an effective method for nanoparticle fabrication; these biogenic nanoparticles could be used as a novel chemotherapeutic drug for the treatment and/or prevention of thyroid cancer.

Disclosure statement

The authors declare that no conflict of interest.

Funding

This work is supported by The Association Foundation Program of Yunnan Provincial Science and Technology, Department and Kunming Medical University under grant Number [2017FE468(-086)]. The author would like to extend their sincere appreciation to the Deanship of Scientific Research at King Saud University for its funding of this research through the Research Group Project No. RGP-300.

References

[1] Gurunathan S, Park JH, Han JW, et al. Comparative assessment of the apoptotic potential of silver nanoparticles synthesized by *Bacillus tequilensis* and *Calocybe indica* in MDA-MB-231 human breast cancer cells: targeting p53 for anticancer therapy. Int J Nanomedicine. 2015;10:4203–4222.

[2] Mukherjee P, Ahmad A, Mandal D. Fungus-mediated synthesis of silver nanoparticles and their immobilization in the mycelial matrix: a novel biological approach to nanoparticle synthesis. Nano Lett. 2001;1(10):515–519.

[3] Conde J, Doria G, Baptista P. Noble metal nanoparticles applications in cancer. J Drug Deliv. 2012;2012:751075.

[4] Ghodake G, Lim SR, Lee DS. Casein hydrolytic peptides mediated green synthesis of antibacterial silver nanoparticles. Colloids Surf B. 2013;108:147–151.

[5] Tanvir S, Oudet F, Pulvin S, et al. Coenzyme based synthesis of silver nanocrystals. Enzyme Microb Technol. 2012;51(4):231–236.

[6] Liu L, Yang J, Xie J, et al. The potent antimicrobial properties of cell penetrating peptide-conjugated silver nanoparticles with excellent selectivity for Gram-positive bacteria over erythrocytes. Nanoscale. 2013;5(9):3834–3840.

[7] Mittal AK, Chisti Y, Banerjee UC. Synthesis of metallic nanoparticles using plant extracts. Biotechnol Adv. 2013;31(2):346–356.

[8] Tuo Y, Liu G, Zhou J, et al. Microbial formation of palladium nanoparticles by *Geobacter sulfurreducens* for chromate reduction. Bioresour Technol. 2013;133:606–611.

[9] Sathyavathi S, Manjula A, Rajendhran J, et al. Extracellular synthesis and characterization of nickel oxide nanoparticles from *Microbacterium* sp. MRS-1 towards bioremediation of nickel electroplating industrial effluent. Bioresour Technol. 2014;165:270–273.

[10] Natarajan K, Selvaraj SV, Murty R. Microbial production of silver Nanoparticles. Dig J Nanomater Biostruct. 2010;5(1):135–140.

[11] Husseiny MI, El-Aziz MA, Badr Y, et al. Biosynthesis of gold nanoparticles using *Pseudomonas aeruginosa*. Spectrochim Acta A . 2007;67(3–4):1003.

[12] Liu L, Shao Z, Ang M, et al. Are microorganisms indispensable in green microbial nanomaterial synthesis. RSC Adv. 2014;4(28):14564.

[13] Zhang X, Yan S, Tyagi RD, et al. Synthesis of nanoparticles by microorganisms and their application in enhancing microbiological reaction rates. Chemosphere. 2011;82(4):489–494.

[14] Pérez-Herrero E, Fernández-Medarde A. Advanced targeted therapies in cancer: drug nanocarriers, the future of chemotherapy. Eur J Pharm Biopharm. 2015;93:52–79.

[15] Hanley C, Layne J, Punnoose A, et al. Preferential killing of cancer cells and activated human T cells using ZnO nanoparticles. Nanotechnology. 2008;19(29):295103.

[16] El-Baghdady K, Elshatoury E, Abdullah MO, et al. Biogenic production of silver nanoparticles by *Enterobacter cloacae* ISM26. Turk J Biol. 2018;42(4):319–321.

[17] Malhotra A, Dolma K, Kaur N, et al. Biosynthesis of gold and silver nanoparticles using a novel marine strain of *Stenotrophomonas*. Bioresour Technol. 2013;31(142):727.

[18] Suresh AK, Pelletier DA, Wang W, et al. Silver nanocrystallites: biofabrication using *Shewanella oneidensis*, and an evaluation of their comparative toxicity on Gram-negative and Gram-positive bacteria. Environ Sci Technol. 2010;44(13):5210–5215.

[19] Buttacavoli M, Albanese NN, Di Cara G, et al. Anticancer activity of biogenerated silver nanoparticles: an integrated proteomic investigation. Oncotarget. 2017;9(11):9685–9705.

[20] Grant CS. Recurrence of papillary thyroid cancer after optimized surgery. Gland Surg. 2015;4(1):52–62.

[21] Davies L, Morris LGT, Haymart M, et al. American association of clinical endocrinologists and American College of Endocrinology Disease State Clinical Review: the increasing incidence of thyroid cancer. Endocr Pract. 2015;21(6):686–696.

[22] Chow S-M, Law SCK, Chan JKC, et al. Papillary microcarcinoma of the thyroid-prognostic significance of lymph node metastasis and multifocality. Cancer. 2003;98(1):31–40.
[23] Gunaseelan S, Balupillai A, Govindasamy K, et al. Linalool prevents oxidative stress activated protein kinases in single UVB-exposed human skin cells. PLOS One. 2017;12(5):e0176699.
[24] Okhawa H, Ohishi N, Yagi K. Assay of lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal Biochem. 1979;95: 351–358.
[25] Kakkar ZYP, Das B, Viswanathan PN. A modified spectrophotometric assay of superoxide dismutase (SOD). Indian J Biochem Biophys. 1984;21:130–132.
[26] Sinha KA. Colorimetric assay of catalase. Anal Biochem. 1972; 47(2):389–394.
[27] Moron MS, De Pierre JW, Mannervik B. Levels of glutathione, glutathione reductase and glutathione S-transferase in rat lung and liver. Biochim Biophys Acta. 1972;82:67–70.
[28] Jesudason E, Masilamoni EG, Charles Jebaraj EJ, et al. Efficacy of DL-alpha lipoic acid against systemic inflammation induced mice: antioxidant defense system. Mol Cell Biochem. 2008;313(1–2): 113–123.
[29] Johnson LV, Walsh ML, Chen LB. Localization of mitochondria in living cells with rhodamine 123. Proc Natl Acad Sci USA. 1980; 77(2):990–994.
[30] Karthikeyan S, Kanimozhi G, Prasad NR, et al. Radiosensitizing effect of ferulic acid on human cervical carcinoma cells *in vitro*. Toxicol In Vitro. 2011;25(7):1366–1375.
[31] Duran N, Marcarto PD, DeSouza GIH, et al. Antibacterial effect of silver nanoparticles produced by fungal process on textile fabrics and their effluent treatment. J Biomed Nanotechnol. 2007;3: 203–208.
[32] Das SK, Khan MR, Guha AK, et al. Silver-nanobiohybrid material: synthesis, characterization and application in water purification. Bioresour Technol. 2012;124:495–499.
[33] Park TJ, Lee SY, Heo NS, et al. In vivo synthesis of diverse metal nanoparticles by recombinant *Escherichia coli*. Angew Chem Int Ed. 2010;49:7019–7024.
[34] Kaler A, Jain S, Banerjee UC. Green and rapid synthesis of anticancerous silver nanoparticles by *Saccharomyces boulardii* and insight into mechanism of nanoparticle synthesis. BioMed Res Int. 2013; 2013:1.
[35] Abo-State MMA, Partila AM. Microbial production of silver nanoparticles by *Pseudomonas aeruginosa* cell free extract. J Ecol Health Environ. 2015; 3(3):91–98.
[36] Kalimuthu K, Suresh Babu R, Venkataraman D, et al. Biosynthesis of silver nanocrystals by *Bacillus licheniformis*. Colloids Surf B. 2008;66:150–153.
[37] Thakkar KN, Mhatre SS, Parikh RY. Biological synthesis of metallic nanoparticles. Nanomedicine. 2010;6(2):257–262.
[38] Zaheer Z, Rafiuddin. Silver nanoparticles to self-assembled films: green synthesis and characterization. Colloids Surf B: Biointerfaces. 2012;90:48–52.
[39] Singh PK, Bhardwaj K, Dubey P, et al. UV-assisted size sampling and anti-bacterial screening of *Lantana camara* leaf extract synthesized silver nanoparticles. RSC Adv. 2015;5(31):24513–24520.
[40] Kumar P, Senthamilselvi S, Lakshmiprabha A, et al. Synthesis of silver nanoparticles from Sargassum tenerrimum and screening phytochemicals for its anti-bacterial activity. Nano Biomed. Eng. 2012;4(1):12–16.
[41] Mosmann T. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods. 1983;65(1–2):55–63.
[42] Farah MA, Ali MA, Chen SM, et al. Silver nanoparticles synthesized from *Adenium obesum* leaf extract induced DNA damage, apoptosis and autophagy via generation of reactive oxygen species. Colloids Surf B: Biointerfaces. 2016;141:158–169.
[43] Behravan M, Hossein Panahi A, Naghizadeh A, et al. Facile green synthesis of silver nanoparticles using *Berberis vulgaris* leaf and root aqueous extract and its antibacterial activity. Int J Biol Macromol. 2019;124:148–154.
[44] Sankar R, Karthik A, Prabu A, et al. *Origanum vulgare* mediated biosynthesis of silver nanoparticles for its antibacterial and anticancer activity. Colloids Surf B Biointerfaces. 2013;1(108):80–84.
[45] Hussain SM, Hess KL, Gearhart JM, et al. In-vitro toxicity of nanoparticles in BRL 3A rat liver cells. Toxicol In Vitro. 2005;19(7): 975–983.
[46] Lee YS, Kim DW, Lee YH, et al. Silver nanoparticles induce apoptosis and G2/M arrest via PKCζ-dependent signaling in A549 lung cells. Arch Toxicol. 2011;85:1529–1540.
[47] AshaRani PV, Low Kah Mun G, Hande MP, et al. Cytotoxicity and genotoxicity of silver nanoparticles in human cells. ACS Nano. 2009;3(2):279–290.

ScienceDirect

Journal of Photochemistry and Photobiology B: Biology
Volume 203, January 2020, 111773

Apoptotic induction and anti-metastatic activity of eugenol encapsulated chitosan nanopolymer on rat glioma C6 cells via alleviating the MMP signaling pathway

Zhenjiang Li [a], Vishnu Priya Veeraraghavan [b], Surapaneni Krishna Mohan [c], Srinivasa Rao Bolla [d], Hariprasath Lakshmanan [e], Subramanian Kumaran [f], Wilson Aruni [g], Aref Ali Mohammed Aladresi [h], Omar H.M. Shair [h], Sulaiman Ali Alharbi [h], Arunachalam Chinnathambi [h]

[a] Department of Neurosurgery, Huaihe Hospital of Henan University, Kaifeng City, Henan Province 475000, China
[b] Department of Biochemistry, Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences, Chennai 600 077, India
[c] Department of Biochemistry, Panimalar Medical College Hospital & Research Institute, Varadharajapuram, Poonamallee, Chennai 600 123, India
[d] Department of Biomedical Sciences, Nazarbayev University School of Medicine, Nazarbayev University, Nur-Sultan City 010000, Kazakhstan
[e] Department of Biochemistry, Karpagam Academy of Higher Education, Coimbatore 641021, Tamilnadu, India
[f] Centre for Drug Discovery and Development, Col Dr.Jeppiaar Research Park, Sathyabama Institute of Science and Technology, Chennai 600119, Tamilnadu, India
[g] School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology, Chennai 600119, Tamilnadu, India
[h] Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

Received 30 November 2019, Revised 24 December 2019, Accepted 31 December 2019, Available online 2 January 2020, Version of Record 10 January 2020.

Check for updates

Show less

Outline | Share | Cite

https://doi.org/10.1016/j.jphotobiol.2019.111773

Get rights and content

Highlights

- Gliomas are one of the most recurrent types of primary intracerebral cancer which are tending to be exceedingly fatal.
- Eu—Cs nanopolymer was assessed by evaluating the protein expression epithelial mesenchymal transition inducer protein.
- Eu—Cs nanopolymers inhibits the expression of NF-κB & EMT protein thereby induces apoptosis.

Full Text
Help

Abstract

Glioma is the prime cause of cancer allied mortality in adolescent people and it accounts about 80% of all malignant tumours. Eugenol is a major bioactive constituent present in the essential oils with numerous pharmacological benefits including nueroprotective activity. The major drawback of eugenol is its extreme volatile property and oxygen sensitivity therefore we increased the efficacy of drug; eugenol by encapsulating with chitosan polymer. Eugenol loaded chitosan polymer (Eu—Cs) was characterized using FTIR, XRD, SEM, HR-TEM analysis and the encapsulation, drug release efficacy was assessed at in vitro

condition. The induction of autophagy and anticancer efficacy of Eu—Cs on glioma cells was evaluated with rat C6 glioma cells using MTT assay, acridine orange staining, immunocytochemical analysis of NFκβ protein expression and FLOW cytometric analysis. The anti-metastatic property of Eu-CS was assessed by immunoblotting and RT-PCR analysis of epithelial mesenchymal transition protein expression in Eu—Cs treated rat C6 glioma cells. Our characterization analysis proves that Eu—Cs possess essential physical and functional properties of copolymer to be utilized as a drug. Further the MTT analysis and AO staining confirms even in the presence of oncogenic inducer and autophagic inhibitors, Eu—Cs exhibits apoptotic potency on rat C6 glioma cells. The result of immunocytochemical studies depicts the inhibition of NFκβ protein expression and flow cytometry studies confirm apoptosis induction by Eu—Cs. The inhibition of metastasis by Eu—Cs was proven by the decrease in epithelial mesenchymal transition protein expression in

Eu-Cs treated rat C6 glioma cells. Over all our results authentically confirms eugenol loaded chitosan nanopolymer persuasively induces apoptosis and inhibits metastasis in rat C6 glioma cells.

Previous Next

Keywords

Glioma; Anticancer drug; Eugenol; Chitosan nanopolymer; C6 glioma cells; Metastasis

Recommended articles

Cited by (67)

Polymeric nanoparticles for drug delivery in glioblastoma: State of the art and future perspectives

2022, Journal of Controlled Release

Show abstract

Chitosan-based biomaterials for the treatment of bone disorders

2022, International Journal of Biological Macromolecules

Show abstract

Chitosan applications in studying and managing osteosarcoma

2021, International Journal of Biological Macromolecules

Show abstract

The effects of chitosan-based materials on glioma: Recent advances in its applications for diagnosis and treatment

2021, International Journal of Biological Macromolecules

Show abstract

Fabrication and assessment of potent anticancer nanoconjugates from chitosan nanoparticles, curcumin, and eug

2022, Frontiers in Bioengineering and Biotechnology

Comparative study of male and female human hair: A microscopic analysis

2022, Journal of Advanced Pharmaceutical Technology and Research

View all citing articles on Scopus

View full text

© 2020 Elsevier B.V. All rights reserved.

Full Text
Help



Copyright © 2023 Elsevier B.V. or its licensors or contributors.
ScienceDirect® is a registered trademark of Elsevier B.V.

RELX™

Full Text
Help

Biotechnological Communication

Biosc.Biotech.Res.Comm. Vol 13 (2) April-May-June 2020 Pp-565-570

BBRC Bioscience Biotechnology Research Communications

Antiseptic Efficacy of Combination of Certain Essential Oils Against Multi Drug Resistant Pathogens

Jerrine Joseph[1], Devadarshini Rajesh[2], Arumugam Suresh[*3], Manikkam Radhakrishnan[1], Thirunvaukarasu Rajasekar[1], Venugopal Gopikrishnan[1], Shakila Suresh[2] and Wilson Aruni[4,5,6]

[1]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai Tamil Nadu, India.

[2]Lingam Microbiological Laboratory, Kanchipuram, Tamil Nadu, India;

[3]Central Research Laboratory Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research Institute, Kanchipuram, Tamil Nadu, India

[4]Sathyabama Institute of Science & -Technology, Chennai – 600119

[5]California University of Science & Medicine, School of Medicine, CA, USA

[6]Musculoskeletal Disease Research Laboratory US Department of Veteran Affairs, Loma Linda, CA, USA.

ABSTRACT

Incorporation of chemicals in the antiseptics and cosmetic products has increased. Natural antiseptics are non-toxic and ideal for human skin; they are eco-friendly and biodegradable. Numerous researches are focusing on herbal plant-based studies for biomedical therapeutic application. This study was designed to examine antiseptic potential of certain essential oils (Lemon grass oil, Cinnamon oil, Citronella, Cedar wood oil, Lemon oil, Lavender oil, Tea tree oil, Sandal wood oil, Basil oil. Essential oils (n = 9) were selected based on the previous studies, the selected essential oils were purchased from Naturoman Pvt. Ltd., Antimicrobial testing of the essential oils was done by CLSI standard, further the antiseptic application of essential oil was tested by American Society for Testing and Materials (ASTM) standards. In addition, Drug resistant clinical pathogens were used to screen antiseptic activity. The following drug resistant pathogens were isolated from our previous studies, Such as *Staphylococcus aureus, Escherichia coli, Klebsiella pneumoniae* and *Candida albicans*. Antioxidant property of essential oil by 2,2-diphenyl-1-picrylhydrazyl (DPPH). 9 essential oils were subjected to antimicrobial testing with drug resistant isolates from DFU, Cinnamon oil, Lavender oil, and Tea tree oil have broad-spectrum activity of 0.125-2µg/ml against drug-resistant *S. aureus E. coli, and C. albicans*. And it has 0.6-0.8µg/ml of antioxidant property done by DPPH method. ASTM standard such as time-kill assay, Anti-biofilm, and Real-life disinfectant efficacy testing showed promising activity for Tea tree oil and Cinnamon amalgamation. The present study showed the importance of essential oil alone and mixture in antiseptic and cosmetic applications. Thus, essential oil blend might be a prospective source of alternative antimicrobial agents and may play a vital role in antiseptic and skin care product development.

KEY WORDS: ANTISEPTIC, ESSENTIAL OILS, MULTIDRUG RESISTANT PATHOGENS.

Article Information:*Corresponding Author: *kpm_suresh@yahoo.com*
Received 18/04/2020 Accepted after revision 16/06/2020
Published: 30th June 2020 Pp-556-564
This is an open access article under Creative Commons License,.
Published by Society for Science & Nature, Bhopal India.
Available at: https://bbrc.in/
Article DOI: http://dx.doi.org/10.21786/bbrc/13.2/30

INTRODUCTION

Antibiotic resistance is documented to be a serious problem that affects the choice of appropriate antibiotic therapy and increases the probability of unfavorable infection outcome. One of the proposed methods to cope with multidrug-resistant (MDR) bacteria is the use of alternative antibacterial treatments, which include natural antimicrobial substances such as plant essential oils (EOs) (Kon et al., 2012). In addition to dramatic events of this type is the ongoing struggle to identify new and effective antibiotics, especially against Gram-negative bacteria; new classes of effective antibiotics have not yet emerged and the prospects of this happening in the near future are low (Piddock et al.,). Clinicians frequently reach a therapeutic limit when treating patients with serious bacterial infections and efforts are being made worldwide to find new treatments for resistant bacteria. New approaches include educating the public, improving sanitation infrastructures, imposing strict regulations on antibiotic prescription, developing new antibiotics, reevaluating old and rejected antibiotics, abolishing the nontherapeutic use of antibiotics (e.g., in agriculture and animal feed), and developing nonantibiotic approaches that can successfully prevent and protect against infectious diseases, (Carlet et al., 2011; Elcocks (2020).

Herbal plants are considered as good antioxidant since ancient times. Aromatic and medicinal plants, through their secondary metabolism, provide a complex mixture of volatile molecules known as essential oils. These volatile molecules exert antibacterial activity that has been used in folk medicine for centuries. During the last few decades, the emergence of antibacterial resistance has forced us to search for new and efficient antimicrobial agents. Moreover, the use of essential oils and their components in combination with antibiotics may increase bacterial susceptibility, thus limiting resistance (Faleiro and Miguel 2013). Essential oils are known to possess potential as natural agents for food preservation. Many of them recently have been qualified as natural antioxidants and proposed as potential substitutes for synthetic antioxidants in specific sectors of food preservation where their use is not in contrast with their aroma (Ruberto & Baratta, 2000). An antioxidant can be broadly defined as any substance that delays or inhibits oxidative damage to a target molecule (Yamagishi et al., 2011). The main characteristic of an antioxidant is its ability to trap free radicals. Antioxidant compounds like phenolic acids, polyphenols and flavonoids scavenge free radicals such as peroxide, hydroperoxide or lipid peroxyl and thus inhibit the oxidative mechanisms that lead to degenerative diseases, (Wu et al., 2011 Elcocks (2020).

The different antioxidant assays is available and, because results rely on different mechanisms, they strictly depend on the antioxidant models employed and on lipophilic/hydrophilic balance. A single-substance/ single-assay produces relative results, and it is perceived as a reductive approach whenever a phytocomplex is involved. Therefore, a multiple-test and a simultaneous chemical characterization must be taken into account whenever assays of essential oils are performed to allow a balance between the sensory acceptability and functional properties (Saccheti et al., 2005).

Hence this present study was undertaken to evaluate the antiseptic efficacy of essential oils

MATERIAL AND METHODS

Essential oils: The following nine essential oils (Lemon grass oil, Cinnamon oil, Citronella, Cedar wood oil, Lemon oil, Lavender oil, Tea tree oil, Sandal wood oil, Basil oil) Purchased from Naturoman Pvt. Ltd(table-1).

Table 1. Selected Essential oil for antimicrobial studies

s.no.	Common name	Botanical Name	Part used
1	Lemon	*Citrus limonum*	Fruit skin
2	Basil	*Ocimum basilicum*	Leaf
3	Cinnamon Bark	*Cinnamomum zeylanicum*	Bark
4	Lavender	*Lavandula officinalis*	Flower
5	Citronella	*Cymbopogon nardus*	Grass
6	Pepper Mint	*Mentha piperita*	Leaf
7	Lemon grass	*Cymbopogon citratus*	Grass
8	Cedar wood	*Cedrus libani*	Wood
9	Eucalyptus	*Eucalyptus polybractea*	Leaf

Antimicrobial Susceptibility test: Preliminary Screening of essential oils for antimicrobial characterization was done by the agar well diffusion method. All the test pathogen with appropriate standard cultures (Escherichia coli ATCC 25922, *Staphylococcus aureus* ATCC 25923, *Enterococcus faecalis* ATCC 29212, *Pseudomonas aeruginosa* ATCC 27853, *Escherichia coli*, ATCC 35218, *Klebsiella pneumoniae* ATCC 700603 and *Candida albicans* ATCC 90028), were prepared according to McFarland standard. The 18 h bacterial cultures were adjusted to approximately 105CFU/ml with sterile saline solution. Lawn culture were made on Mueller-Hinton agar using a sterile cotton swab for uniform microbial growth. The essential oils were diluted in DMSO with 0.5% Tween 80. Well were made at 6mm diameter in inoculated MHA agar plate. A standard mineral oil and 5µg concentration of Ciprofloxacin was used as positive control. Then the plates were incubated at 37°C for overnight. After incubation, the zone of inhibition was measured with standard procedures.

Minimum Inhibitory concentration were performed 9 oils based on preliminary antimicrobial results such as Cinnamon, lemon, lemon grass, and tea tree oil were identified to have potent antimicrobial activity and their Minimum Inhibitory Concentrations (MIC) were determined. The standard broth dilution method followed by the Clinical Laboratory standard institute, 2018. Standard two fold dilution method followed for

each oil, ranging from 1 to 512 µg/ml, was prepared in Muellur Hinton broth. 0.5 McFarland standard broth culture suspension were prepared and inoculated in individual MIC tube, Inoculated tubes were incubated at 37°C for overnight for bacteria for fungi 24-72 hrs based. The MIC were determined as the lowest concentration of oil inhibiting visible growth of each organism on the tube.

Organisms used: The bacterial test strains used in this study were Staphylococcus aureus ATCC 29213, *Escherichia coli* ATCC 25922, *Pseudomonas aeruginosa* ATCC 27853, *Enterococcus faecalis* ATCC 29212, *Candida krusei* ATCC 6258, *Candida albicans* ATTC 10231. clinical pathogens were used to screen antiseptic efficacy, the pathogens were isolated from my previous studies, it was stored and maintained, Drug resistant *S. aureus* (31), *E. coli* (11), *K. pneumoniae* (07) and *C. albicans* (08). The bacterial strains were maintained on 60% glycerol BHI broth at -200C in the Lingam Microbiological Laboratory, Kanchipuram.

Synergistic effect of Essential oil: Effect of Tea tree oil and cinnamon oil; Tea tree oil and Lemon grass oil mixture were tested for synergistic activity against routine bacterial, viral and fungal pathogens. This study could serve as a baseline data to investigate new combination of essential oil for antiseptics.

Antioxidant activity of essential oil by DPPH method: The hydrogen atoms or electrons donation ability of the corresponding extracts and some pure compounds were measured by bleaching the purple colored methanolic solution of DPPH. The effects of methanolic extract and essential oil on DPPH radicals were evaluated according to a method described elsewhere (21). 4 mL samples of various concentrations of the extracts in methanol were separately added to a 1 mL solution of DPPH radical in methanol (final concentration of DPPH was 0.2 mM). The mixture was shaken vigorously and allowed to stand for 30 min after which the absorbance of the resulting solution was measured at 517 nm with a spectrophotometer.

Inhibition of free radical DPPH as percentage [I(%)] was calculated as follows:

$$I\ (\%) = 100 \times (A\text{blank} - A\text{ sample}) / A\text{blank}$$

where A blank is the absorbance of the control (containing all reagents except the test compound) and A sample is the absorbance of the test compound. EC50 value (µg /mL) is the effective concentration at which DPPH radicals are scavenged by 50%. This was obtained by interpolation and using linear regression analysis.

Time kill assay for the essential oil blend: Sterile 96 well microtitre plates were used test with Muller Hinton broth for the determination of the time kill assay. Briefly, 180 µl of sterile freshly prepared MH broth was added in each well. 10µl of essential oil were added to the media. 10 µl of the standard bacterial strains was inoculated in test medium. Every 30mins, killing activity was measured by spot culture in new plate MHA, this procedure repated every 20mins for 6 hrs.

Anti-biofilm assays: Antibiofilm activity was carried out in polystyrene microtitre plate method. An overnight culture of standard strains was diluted to the McFarland standard 0.5. After the addition of different concentrations of Essential oil1-128 µl/ml of Muller Hinton broth, 10-µl broth culture was inoculated in all tube except negative control. incubated for 48 to 72 h at 37^{0}C. Thereafter Planktonic cells of Standard strains is discarded, 100µl of 1% (w/v) aqueous solution of crystal violet will be added to microtitre plate and incubated at room temperature for 20 minutes. Then crystal violet solution is removed from the microtitre plate and it will be washed thoroughly with sterile water. For quantification of attached cells, the bound crystal violet solubilized in dimethyl sulfoxide (DMSO) and spot culture method.

RESULTS AND DISCUSSION

Antimicrobial assay: Agar well Diffusion method: Lemon grass oil showed highest antimicrobial activities against *C. krusei*_ATCC 6258 (19 mm) followed by *E. faecalis*_ATCC 29212 (18mm), *E. coli*_ATCC 25922 and *P. aeruginosa*_ATCC 27853 (17mm), *S. aureus*_ATCC 29213 (13mm) and minimum inhibition against *C. albicans*_ATTC 10231 (12mm).*Cinnamon oil* showed highest antimicrobial activities against *S. aureus*_ATCC 29213 (30mm), *E. coli*_ATCC 25922 and *E. faecalis*_ATCC 29212 (25mm), *C. albicans*_ATTC 10231 (20mm), *C. krusei*_ATCC 6258 (18 mm) and minimum inhibition against *P. aeruginosa*_ATCC 27853 (16mm).*Citronella oil* showed highest antimicrobial activities against *E. faecalis*_ATCC 29212 (12mm) followed by *P. aeruginosa*_ ATCC 27853 (11mm) and minimum inhibition against *S. aureus*_ATCC 29213 (7mm). other organisms not shown any activities against other organisms .Cedar wood oil showed highest antimicrobial activities against *E. coli*_ATCC 25922 (11mm) and minimum activities against *S. aureus*_ATCC 29213 (10mm).

Lemon oil showed highest antimicrobial activities against *S. aureus*_ATCC 29213 (24mm), *P. aeruginosa*_ ATCC 27853 (21mm) *E. coli*_ATCC 25922 (20mm), *E. faecalis*_ATCC 29212 (17mm), *C. albicans*_ATTC 10231 (15mm) and minimum inhibition activities against *C. krusei*_ATCC 6258 (13 mm).Lavender showed highest antimicrobial activities against *E. faecalis*_ATCC 29212 (12mm), *P. aeruginosa*_ATCC 27853 and *C. krusei*_ATCC 6258 (11 mm), *E. coli*_ATCC 25922 and *C. albicans*_ATTC 10231 (10mm) and minimum inhibition activities against *S. aureus*_ATCC 29213 (9mm).Tea tree oil showed highest antimicrobial activities against *S. aureus*_ATCC 29213 (35mm), followed by *C. albicans*_ATTC 10231 (32mm), *C. krusei*_ATCC 6258 and *E. coli*_ATCC 25922 (28mm) *E. faecalis*_ATCC 29212 (26mm) and minimum inhibition activities against *P. aeruginosa*_ATCC 27853 (25mm). Sandal wood oil showed highest antimicrobial activities against *E. coli*_ATCC 25922 (11mm), *E. faecalis*_ATCC 29212 (10mm) and minimum inhibition activities against

*S. aureus*_ATCC 29213 (7mm).Basil oil showed highest antimicrobial activities against *S. aureus*_ATCC 29213 (11mm), *C. krusei*_ATCC 6258 (10 mm) and minimum inhibition activities against *P. aeruginosa*_ATCC 27853 (9mm) (table-2)

Table 2. Antimicrobial property of essential oils

Microbes	Agar well diffusion assay									Minimum Inhibitory Concentration(µL)			
	Lemon grass	Cinnamon oil	Citronella oil	Cedar wood	Lemon oil oil	Lavender	Tea tree	Sandal wood oil	Basil oil oil	Lemon grass	Cinnamon oil	Lemon oil	Tea tree oil
S. aureus ATCC 29213	13	30	7	10	24	9	35	7	11	64	1	32	0.5
E. coli ATCC 25922	17	25	-	11	20	10	28	11	-	16	1	64	1
P. aeruginosa ATCC 27853	17	16	11	-	21	11	25	-	9	64	2	128	1
E. faecalis ATCC 29212	18	25	12	-	17	12	26	10	-	32	1	64	0.12
C. krusei ATCC 6258	19	18	-	-	13	11	28	-	10	-	-	-	-
C. albicans ATTC 10231 Multi drug Resistant diabetic foot ulcer isolates	12	20	-	-	15	10	32	-	-	16	4	128	1
MRSA	11	32	-	-	18	7	28	-	10	128	0.5	64	0.5
P. aeruginosa	14	30	6	-	21	8	28	15	-	64	1	64	2
E. coli	13	20	11	-	20	10	25	-	9	32	0.5	128	0.5
K. pneumoniae	12	22	-	8	14	8	26	-	-	64	1	128	0.25
C. albicans	14	19	-	-	13	9	28	9	10	32	8	64	1

Effect of essential oil with Multi drug Resistant isolates : Lemon grass oil showed highest antimicrobial activities against *P. aeruginosa* and *C. albicans* (14mm) followed by *E. coli* showed (13mm), *K. pneumonia* (12mm) and minimum inhibition activities against MRSA (11mm). Cinnamon oil showed highest antimicrobial activities MRSA (32m), followed by *P. aeruginosa* (30mm), *K. pneumonia* (22mm), *E. coli* (20mm) and minimum inhibition activities against *C. albicans* (19mm). *Citronella oil* showed highest antimicrobial activities *E. coli* (11mm) and minimum inhibition activities against *P. aeruginosa* (6mm).Cedar wood oil showed highest antimicrobial activities *K. pneumonia* (8mm). Other organisms not shown any activities. Lemon oil showed highest antimicrobial activities *P. aeruginosa* (21mm) followed by *E. coli* (20mm), MRSA (18mm), *K. pneumonia* (14mm) and minimum inhibition activities against *C. albicans* (13mm).Lavender oil showed highest antimicrobial activities *E. coli* (10mm) followed by *C. albicans* (9mm), *P. aeruginosa* and *K. pneumonia* (8mm) and minimum inhibition activities against MRSA (7mm).Tea tree oil showed highest antimicrobial activities MRSA, *P. aeruginosa* and *C. albicans* (28mm), followed by *K. pneumonia* (26mm), minimum inhibition activities against *E. coli* (25mm).Sandal wood oil showed highest antimicrobial activities *P. aeruginosa* (15mm) and minimum inhibition activities against *C. albicans* (9mm).Basil oil showed highest antimicrobial activities MRSA and *C. albicans* (10mm) and minimum inhibition activities against *E. coli* (9mm)

Table-3. Anti-biofilm activity of essential oil by viability

Microbes	Essential oil (µL/mL)			
	Lemon grass oil	Cinnamon oil	Lemon oil	Tea tree oil
S. aureus ATCC 29213	16	2	128	4
E. coli ATCC 25922	16	2	128	2
P. aeruginosa ATCC 27853	128	2	128	4
E. faecalis ATCC 29212	64	2	32	2
C. krusei ATCC 6258	-	-	-	-
C. albicans ATTC 10231	64	8	128	4

Oil Blend: Cinnamon oil: Tea tree oil=1:0.25; 1:0.5; 1:1; 0.5:1; and 1:0.25

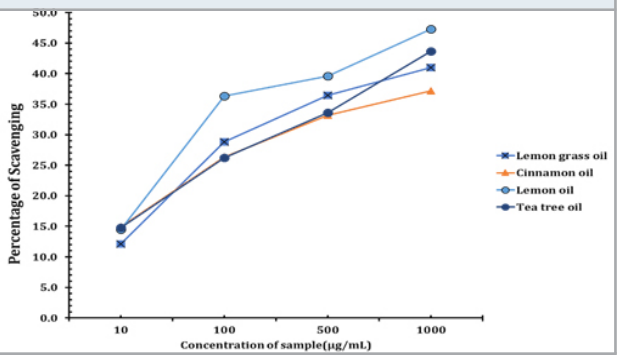
Minimum Inhibitory Concentration: Based on the antimicrobial activities Lemon grass oil, Cinnamon oil, Lemon oil and Tea tree oil taken for the Minimum

Inhibitory Concentration. Lemon grass oil showed maximum Minimum Inhibitory Concentration values against *S. aureus*_ATCC 29213 and *P. aeruginosa*_ATCC 27853 (64 µL) followed by *E. faecalis*_ATCC 29212 (32 µL), *E. coli*_ATCC 25922 and *C. albicans*_ATTC 10231 (16 µL).Cinnamon oil showed maximum Minimum Inhibitory Concentration values against *C. albicans*_ATTC 10231 (4µL), *P. aeruginosa*_ATCC 27853 (2µL), *S. aureus*_ATCC 29213, *E. coli*_ATCC 25922 and *E. faecalis*_ATCC 29212 (1µL),Lemon oil showed maximum Minimum Inhibitory Concentration values against *P. aeruginosa*_ATCC 27853 and *C. albicans*_ATTC 10231 (128 µL), *E. coli*_ATCC 25922 and *E. faecalis*_ATCC 29212 (64µL), *S. aureus*_ATCC 29213 (32 µL).Tea tree oil showed maximum Minimum Inhibitory Concentration values *E. coli*_ATCC 25922, *P. aeruginosa*_ATCC 27853 and *C. albicans*_ATTC 10231 (1 µL), *S. aureus*_ATCC 29213 (0.5 µL), *E. faecalis*_ATCC 29212 (0.12µL) (table-2)

Table 4. Minimum Inhibitory concentration of essential oil Blend-

	Tea tree oil:	cinnamon oil	Essential oil	blend	(µL/mL)
Microbes	1:0.25	1:0.5	1:1	0.5:1	0.25:1
S. aureus ATCC 29213	16	1	1	2	2
E. coli ATCC 25922	16	0.5	0.5	1	0.5
P. aeruginosa ATCC 27853	128	1	0.5	4	2
E. faecalis ATCC 29212	64	2	0.12	0.5	1
C. krusei ATCC 6258	-	-	0.5	2	2
C. albicans ATTC 10231	64	4	0.25	2	4

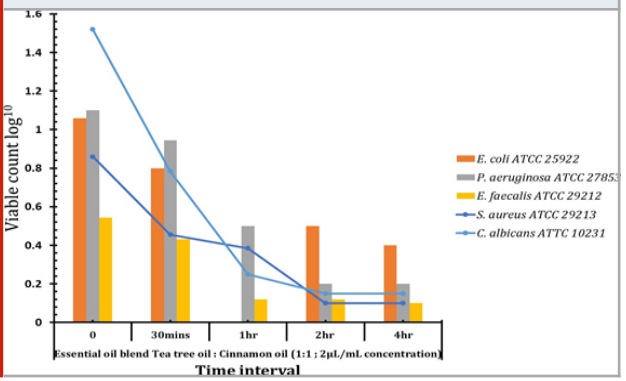
Figure 1: Antioxidant activity of essential oil by DPPH method



Effect of essential oil with Multidrug resistant isolates Based on the antimicrobial activities of Multi drug Resistant diabetic foot ulcer isolates.

Lemon grass oil, MRSA (128 µL) *P. aeruginosa* and *K. pneumonia* (64 µL) , *E. coli* and *C. albicans* (32 µL)) Cinnamon oil *C. albicans* (8µL), *P. aeruginosa* and *K. pneumonia* (1µL), MRSA and *E. coli* (0.5 µL) Lemon oil- *E. coli* and *K. pneumonia* (128µL), MRSA, *C. albicans* and *P. aeruginosa* (64 µL) Tea tree oil- *P. aeruginosa* (2 µL), *C. albicans* (1 µL) MRSA and *E. coli* (0.5 µL), *K. pneumonia* (0.25 µL)

Figure 2: Time kill assay for the essential oil blend



Antioxidant and Time kill assay: Antioxidant of essential oil was shown in figure 1. Lemon oil showed Highest percentage of anti-oxidant activities when compare with other oil. Time kill assay for the essential oil blend show in figure 2. Better growth suppression or killing assay was started on 30 mins above at the concentration of 2 µL of 1:1 Cinnamon oil and Tea tree oils. Recent study found cinnamon oil alone exhibited a broad-spectrum activity against Gram-negative and Gram-positive bacteria and showed bacteriostatic and bactericidal effects against *P. aeruginosa* PAO1 at 2%, including known multidrug resistant species. The number of scientific studies on the antimicrobial activity of plant essential oils has strongly grown over the last three decades. Due to the use of many different microbiological methods for susceptibility testing and different definitions of antimicrobial activity, the comparability of studies on essential oils is often critical. Many studies focus on selected EOs, providing insight into their activity against one or more microorganisms, but only few publications compress information by testing multiplicities of essential oils with a defined single method. Therefore, this study was conducted to investigate as many EOs as possible with a single quantitative microbiological method, to achieve maximum comparability.

CONCLUSION

The present study showed the importance of essential oil alone and mixture in antiseptic and cosmetic applications. Thus, essential oil blend might be a prospective source of alternative antimicrobial agents and may play a vital role in antiseptic and skin care product development.

Authors Contributions: All authors have equal contribution in bringing out this research work.

Conflict of Interest: None.

REFERENCES

Carlet J, Jarlier V, Harbath S, Voss A, Goossens H, Pittet D.(2012) Ready for a world without antibiotics? The Pensieres antibiotic resistance call to action. Antimicrob Res Inf Cont 2012;1:11.

Elcocks ER (2020) Spencer-Phillips PTN, Adukwu EC. Rapid bactericidal effect of cinnamon bark essential oil against *Pseudomonas aeruginosa*. J App Microbiol 128(4): 1025-1037

Faleiro ML, Miguel MG.(2013) Use of essential oils and their components against multidrug-resistant bacteria. In Fighting multidrug resistance with herbal extracts, essential oils and their components 2013 Jan 1 (pp. 65-94). Academic Press.

Kon KV, Rai MK.(2012) Plant essential oils and their constituents in coping with multidrug-resistant bacteria. Expert review of anti-infective therapy Jul 1;10(7):775-90.

Piddock LJ. (2011)The crisis of no new antibiotics-what is the way forward? Lancet Infect. Dis 12:249-53

Ruberto G, Baratta MT (2000): Antioxidant activity of selected essential oil components in two lipid model systems. Food Chem 69: 167–174.

Saccheti G, Maietti S, Muzzoli M, Scaglianti M, Manfredini S, Radice M, Bruni R (2005): Comparative evaluation of 11 essential oils of different origin as functional antioxidants, antiradicals and antimicrobials in foods. Food Chem 91: 621–632.

Wu YY, Li W, Xu Y, Jin EH, Tu YY.(2011) Evaluation of the antioxidant effects of four main theaflavin derivatives through chemiluminescence and DNA damage analyses. J Zhejiang Univ Sci B. 12: 744–751.

Yamagishi S, Matsui T.(2011) Nitric oxide, a Janus-faced therapeutic target for diabetic microangiopathy-Friend or foe? Pharmacol Res. 64:187–194.

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

Original Article

AntiMycobacterial Activity of Endophytic Actinobacteria from Selected Medicinal Plants

Alka Ashok[1], Usha Nandhini[1], Ayswarya Sreenivasan[2], Manigundan Kaari[2], Revathy Kalyanasundaram[2], Radhakrishnan Manikkam[2]

[1]Department of Biotechnology, Sathyabama Institute of Science and Technology, [2]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

Abstract

Background: Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (MTB), poses a serious threat to human life. In addition, certain nontuberculous mycobacterial (NTM) species also cause infections in immune suppressed and even in immunocompetent individuals. With this view, the present study investigated the antimycobacterial properties of endophytic actinobacteria using *Mycobacterium smegmatis* as a surrogate model. **Methods:** Endophytic actinobacterial cultures were isolated from seven medicinal plants and their antimicrobial activity against two Gram-positive bacteria, three Gram-negative bacteria, fungi, and *M. smegmatis* was studied. Crude bioactive metabolites from one potential culture EAM4 were tested against *Staphylococcus aureus, Escherichia coli*, and *M. smegmatis* by disc-diffusion method, whereas activity against MTB H37Rv was tested by luciferase reporter phage (LRP) assay. The potential actinobacterial culture EAM4 was identified based on their 16S rRNA analysis. **Results:** Thirty actinobacterial cultures were isolated and were belonged to the genus *Streptomyces*. Maximum number of actinobacterial strains was isolated from *Selaginella stellata* (36.6%) and *Pilea microphylla* (36.6%). Gram-positive pathogens are inhibited by more number of actinobacterial cultures (>90%) than Gram-negative pathogens (50%). Twenty-eight of thirty actinobacterial cultures showed inhibitory activity against *M. smegmatis*. Bioactive metabolites from the potential culture EAM4 showed better inhibition against the *S. aureus*, *E. coli*, and *M. smegmatis*. In the LRP assay, 79% inhibition was exhibited against the MTB H37Rv. The potential actinobacterial strain EAM4 was identified as *Streptomyces* sp. based on their 16S rRNA analysis. **Conclusion:** Findings of the present study revealed that the endophytic cultures of *Streptomyces* investigated in this study are the promising source for isolating antimycobacterial metabolites effective against TB and nontuberculous mycobacterial pathogens.

Keywords: Antimycobacterial, endophytic, *Mycobacterium smegmatis, Mycobacterium tuberculosis, Streptomyces*

Introduction

Worldwide, tuberculosis (TB) remains among the top 10 causes of death. According to the World Health Organization Report (2019), 10 million peoples are estimated to fell sick with TB worldwide and 1.5 million people died.[1] Typically, after the 6–8 months of intensive antibiotic treatment, about 80%–90% of drug-sensitive TB patients are cured. However, the side effects due to combination anti-TB therapies result in the nonadherence of TB patients for treatment which, in turn, lead to the emergence of high levels of drug-resistant *M. tuberculosis* (MTB) strains. In addition, it takes up to 24 months of combination chemotherapy for the treatment of drug resistant TB.[2] Hence, there is a need to discover novel drugs to treat TB with lower toxicities and shorter duration of treatment.[3]

In addition, at the global level, the frequency and mortality due to diseases caused by the pathogenic nontuberculous mycobacteria (NTM) have been increasing constantly. In early days, NTMs are thought to be harmless environmental saprophyte bacterial members of mycobacterial genera and only very few species are risky to immune-suppressed individuals or

Access this article online

Quick Response Code:

Website: www.bmbtrj.org

DOI: 10.4103/bbrj.bbrj_73_20

Address for correspondence: Dr. Radhakrishnan Manikkam, Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India.
E-mail: mrkactinos@gmail.com

ORCID: https://orcid.org/0000-0001-7616-8254

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

How to cite this article: Ashok A, Nandhini U, Sreenivasan A, Kaari M, Kalyanasundaram R, Manikkam R. AntiMycobacterial activity of endophytic actinobacteria from selected medicinal plants. Biomed Biotechnol Res J 2020;4:193-9.

Submitted: 21-May-2020; **Revised:** 27-Jun-2020; **Accepted:** 28-Jun-2020; **Published:** 12-Sep-2020

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

those with imperfect lung structure. Nevertheless, now, some NTM species are infecting children and adults at a higher rate, even those who are immunocompetent.[4] Unfortunately, like TB, NTM infections are incredibly difficult to treat due to the inherent resistance of these bacteria to many common antibiotics.[5,6] Hence, there is a timely need to discover new antibiotics effective against MTB and other NTM pathogens.

Microbes, especially the members of the phylum Actinobacteria, are the prolific producers of anti-TB antibiotics for several decades.[7-9] Two important antibiotics, streptomycin and rifamycin, in current anti-TB regimen are produced by the members of Actinobacteria. Several other anti-TB metabolites were also reported from different actinobacterial genera.[10,11] However, there is a turndown in the newness of antimycobacterial metabolites from Actinobacteria due to the comprehensive searching of usual sources which results in the re-discovery of known metabolites. As an alternative, exploring Actinobacteria from underrepresented sources or ecosystems may find the lane for isolating novel antimycobacterial metabolites.

Medicinal plants are customarily explored for several pharmaceutical compounds including those with anti-TB properties. However, none of them reached the marketable level for TB treatment. In parallel, microbes associated with the healthy plant parts without causing any ill effects to their host plants, called endophytes, are documented as the richest source for novel bioactive metabolites with useful antibacterial, antifungal, and anticancer properties.[12,13] There are few anti-TB antibiotics such as munumbicin B effective against multidrug-resistant (MDR) MTB that are reported from endophytic *Streptomyces*.[14] The present study investigated the endophytic actinobacteria from selected medicinal plants collected from Kerala, India, for their antimycobacterial properties against *Mycobacterium smegmatis* and MTB.

Methods

Sample collection and pretreatment

A total of 7 medicinal plants as listed in Table 1 were collected from the Vadakara region of Kerala state in India during November 2018. Specific plant parts, as mentioned in Table 1, from each plant were collected and thoroughly washed for a minute using running tap water to take away the soil and other organic debris. The washed plant parts were dried, crumbled into small fragments, and processed with a four-step surface sterilization method. Briefly, plant parts are separately sterilized by immersing it in 70% ethanol for 1 min, sodium hypochlorite for 5 min, 90% ethanol for 30 s, and a final rinse with distilled water.[15]

Isolation of endophytic actinobacteria

Each sample was then ground using a clean motor and pestle with 5 ml of distilled water. The resultant homogenates were serially diluted 10^{-5} dilutions using sterile distilled water. Hundred microliter aliquots from the 10^{-3}, 10^{-4}, and 10^{-5} dilutions were plated on starch casein nitrate agar (SCA) media (supplemented with nalidixic acid (20 µg/ml and nystatin 50 µg/ml) and spread using flame sterilized L-rod. Colonies with actinobacterial morphology have been selected after 4 weeks of incubation at 28°C.[16] Morphologically, different colonies of Actinobacteria were retrieved from SCA with suspected actinobacterial morphology that have been retrieved from incubated plates and inoculated onto Yeast Extract Malt Extract (YEME) agar medium and incubated for 7 days at 28°C. The retrieved actinobacterial cultures were stored in glycerol broth (20%) and slants of YME agar.[17]

Table 1: Endophytic actinobacteria isolation from medicinal plants

Name of the plant	Plant part used	Number of strains isolated in the present study
Aloe vera	Leaf	0
Sida cordifolia	Root	4
Sauropus androgynus	Leaf	0
Ocimum tenuiflorum	Leaf	1
Piper nigrum	Leaf	2
	Seed	1
Selaginella stellata	Leaf	2
	Root	9
Pilea microphylla	Leaf	8
	Stem	3

Characterization of endophytic actinobacteria

Cultural characteristics of actinobacterial cultures were recorded after 10 days of growth at 28°C on the YEME agar medium. Actinobacterial cultures were observed for their growth level, colony consistency, the aerial mycelia color, reverse side pigment, and soluble pigment production. Micromorphological features were examined by the method of slide culture using YEME agar. On the 10th day of incubation, slides were observed under a bright-field microscope with × 40. The noted micromorphological features include the vegetative mycelium, reproductive mycelium, and mycelia fragmentation if any.[18]

Screening for *in vitro* antimicrobial activity

Agar plug method[17] was followed to screen the actinobacterial cultures against a panel of test pathogens, namely *Staphylococcus aureus* ATCC 29213, *Bacillus cereus*, *Escherichia coli* ATCC 25922, *Pseudomonas aeruginosa* ATCC 27853, *Klebsiella pneumoniae* ATCC 13882, and *Candida albicans*. All the pathogenic cultures were grown on Mueller Hinton broth by incubating at 37°C for 18 h. Inoculum with 0.5 McFarland standard was spread using sterile cotton swan on Mueller Hinton Agar (MHA) plates. For the production of extracellular metabolites, all the endophytic actinobacterial cultures were cultivated for 10 days on YEME agar plates at 28°C. After incubation, the mycelial biomass was collected from the agar plates with the help of a sterile spatula. Then, 5-mm diameter agar plug containing the secreted actinobacterial metabolites was taken from the YEME agar medium and placed over MHA plates. The inhibition zone was measured after 24 h of incubation at 37°C and expressed in millimeter in diameter.

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

Screening for antimycobacterial activity using *Mycobacterium smegmatis* as surrogate

All the actinobacterial strains were tested for antimycobacterial activity against the NTM, *M. smegmatis*, by adopting the agar plug method as described above with some modifications. A loopful of *M. smegmatis* colonies were inoculated into 0.3 ml of Middlebrook 7H9 broth in a sterile Bijou bottle added with glass beads. The content was vortexed for 30 s and kept undisturbed for 5 min. Then, the suspension was made up to 5 ml using Middlebrook 7H9 broth. Then, 200 µl of *M. smegmatis* suspension was added to 5 ml of molten 7H11 agar and poured immediately onto Middlebrook 7H9 agar plate. Agar plug (5 mm in diameter) of actinobacterial cultures from the YEME agar was placed over *M. smegmatis* seeded MHA plates. The plates were sealed and incubated at 37°C for 48 h. The zone of inhibition was measured after incubation.

Bioactive metabolites production from EAM4

Endophytic actinobacterial culture EAM4, which showed maximum inhibition against *M. smegmatis* and other microbial pathogens, was selected for the bioactive metabolites production through agar surface fermentation.[17] Spores of the potential culture EAM4 were inoculated onto five YEME agar plates for the secretion of extracellular secondary metabolites. The mycelial growth from the culture plate was scrapped after 10 days of incubation with the help of a sterile spatula. The remaining agar medium which contains the secreted secondary metabolites was cut into pieces and extracted using ethyl acetate, methanol, chloroform, and hexane separately (1:2 ratio) for 24 h. The solvent portion was collected and concentrated using a rotary evaporator and quantified.

In vitro screening of extracts against *Staphylococcus aureus, Escherichia coli*, and *Mycobacterium smegmatis*

The solvent extracts of the actinobacterial culture EAM4 were tested by adopting the method of well diffusion against *S. aureus, E. coli*, and *M. smegmatis*. The main stock of the ethyl acetate extract at 10 mg/ml concentration was prepared using DMSO. The main stock was further filtered using a syringe filter (0.45 µ). One mg per ml of working stock solution was prepared by adding 100 µl of stock extract into 900 µl of solvent and mixed well. Inoculum of all three test pathogens was prepared and inoculated on to respective agar media by adopting the same procedure described in the preliminary screening. A well with 5 mm in diameter size was made on all the plates using sterile cork borer and loaded with 20 µl of different solvent extracts obtained from the culture EAM4. The inhibition zone was measured after 24–48 h of incubation at 37°C.

In vitro screening of extracts against *Mycobacterium. tuberculosis* H37Rv

The antitubercular activity of the ethyl acetate extract was tested by adopting luciferase reporter phage (LRP) assay[17] using MTB H37Rv as a test pathogen. Two mg per ml of working concentration of sample was prepared using 10% dimethyl sulfoxide (DMSO) and filtered using 0.45 µm filters. The viability of MTB H37Rv culture was maintained on Lowenstein–Jensen medium at 30°C ± 2°C.

High titer of mycobacteriophage phAETRC202 (obtained from the Department of Bacteriology, National Institute for Research in Tuberculosis, Chennai, India) was propagated using *M. smegmatis* in Middlebrook 7H9 complete medium. Approximately 350 µl of G7H9 broth supplemented with 10% albumin dextrose and 0.5% glycerol taken in cryovials and 50 µl of actinobacterial extract from main stock was added into it to achieve a final concentration of 200 µg/ml. Then, 100 µl cell suspension of MTB H37Rv was added into all the vials. The 10% DMSO was also included in the assay as the solvent control. After incubation for 72 h at 37°C, 50 µl of high-titer phage phAETRC202 and 40 µl of 0.1 M CaCl2 solution were added into the test and control vials. All the reaction vials were once again incubated for 4 h at 37°C. After incubation, 100 µl of the reaction mixture from each vial was transferred to luminometer cuvette and added with 100 µl of D-Luciferin solution. The relative light unit (RLU) was measured using the luminometer (Berthold) at 10 s interval time. Extracts showing a 50% reduction or more in RLU were considered as active against MTB.

$$\%\ \text{RLU reduction} = \frac{\text{control RLU} - \text{test RLU}}{\text{control RLU}} \times 100$$

Molecular characterization of potential actinobacterial culture EAM4 by 16S rRNA analysis

The genomic DNA of the potential culture EAM4 was isolated using solute ready genomic DNA kit and further analyzed by gel electrophoresis and quantified using a spectrophotometer (NanoDrop ND-1000, Thermo Scientific, Gloucester, UK). The 16S rRNA gene sequence of strain EAM4 was amplified using the primers: 27F 5′ AGAGTTTGATCMTGGCTCAG3′ (forward) and 1492R 5′ TACGGYTACCTTGTTACGACTT3′ (reverse). The polymerase chain reaction (PCR)-amplified product was sequenced and analyzed at Eurofins Genomics, Bangalore, India. The obtained 16S rRNA gene sequence was aligned with similar sequences available in GenBank using the MEGA 7 program. The aligned sequences were used to construct the phylogenetic tree by following the neighbor-joining algorithm in the MEGA 7 program.[19] The bootstrap method was used to determine the confidence of the branches (phylogenetic tree branches).[20] The partial 16S rRNA nucleotide sequence of strain EAM4 has been deposited into the GenBank database.

Results and Discussion

Isolation and characterization of endophytic actinobacterial strains

In this present study, a total of 30 endophytic actinobacterial cultures (EAM 1–EAM 30) were isolated from seven

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

medicinal plants collected from Kerala, India. The number of actinobacterial cultures isolated from different plants is given in Table 1. In some earlier studies, endophytic actinobacteria are reported from *Aloe vera,*[21] *Ocimum tenuiflorum*,[22,23] and *Piper nigrum*[24] but not from *Sida cordifolia, Sauropus androgynus, Selaginella stellate,* and *Pilea microphylla*. In the present study, the maximum number of actinobacterial cultures was isolated from *S. stellate* (36.6%) and *P. microphylla* (36.6%) [Table 1]. Morphological characteristics of endophytic actinobacterial cultures are given in Table 2. On the YEME agar, the maximum of the cultures (90%) showed good growth with powdery colonies (80%) with grey aerial mycelium (50%). About 46.6% of the cultures produced brown reverse side pigment and 33.3% of the cultures produced brown or pale brown soluble pigments. All the cultures showed the presence of both vegetative (substrate) and reproductive (aerial) mycelium under bright filed microscopic observation. Based on the phenotypic characteristics, the isolated cultures were tentatively identified as *Streptomyces*. Similarly, in previous studies also, *Streptomyces* species are the predominant endophytic actinobacterial genera reported from different medicinal plants.[25,26]

In vitro antimicrobial activity

In the preliminary screening for antimicrobial activity, among the 30 actinobacterial cultures, 27 cultures were active against *S. aureus* and 19 strains were active against *B. cereus*. Among the Gram-negative bacterial pathogens tested, *E. coli* was inhibited by a maximum of 16 actinobacterial cultures, followed by *K. pneumoniae* being inhibited by 15 strains, whereas *P. aeruginosa* was inhibited by 11 actinobacterial cultures [Figure 1a]. Sixteen actinobacterial strains were found to be active against *C. albicans*. Interestingly, actinobacterial cultures such as EAM3, EAM4, EAM6, EAM12, and EAM17 showed broad-spectrum activity against both the Gram-positive and Gram-negative bacterial pathogens and the fungi *C. albicans* [Table 3]. Similarly, *Streptomyces* sp., isolated from *A. vera* was exhibited maximum inhibition against Gram-positive bacteria, *S. aureus* and *Micrococcus luteus* followed by *Bacillus subtilis* and fungal pathogens *Aspergillus niger*, followed by *Candida tropicalis*, *C. albicans*, and *Microsporum canis*.[21] According to the early reports, the maximum number of actinobacteria inhibits Gram-positive bacteria than Gram-negative bacteria.[27] However, in this study, about 50% of the endophytic actinobacterial cultures

Table 2: Morphological characteristics of endophytic actinobacterial strains

Strain number	Growth	Consistency	AMC	RSP	SP	AM	SM
EAM 1	Good	Powdery	Gray	Brown	Pale brown	+	+
EAM 2	Good	Powdery	Gray	Yellow	−	+	+
EAM 3	Moderate	Pasty	White	Brown	Brown	+	+
EAM 4	Good	Powdery	White	Brown	Brown	+	+
EAM 5	Good	Powdery	Gray	Orange	Pale brown	+	+
EAM 6	Good	Powdery	Gray	Brown	Brown	+	+
EAM 7	Good	Powdery	Gray	Orange	−	+	+
EAM 8	Good	Powdery	White	Black	Brown	+	+
EAM 9	Moderate	Pasty	White	Orange	−	+	+
EAM 10	Good	Powdery	Gray	Brown	pale gray	+	+
EAM 11	Good	Leathery	Grayish yellow	Pale yellow	−	+	+
EAM 12	Moderate	Powdery	Light gray	Pale yellow	−	+	+
EAM 13	Good	Powdery	Light gray	Yellow	Pale yellow	+	+
EAM 14	Good	Powdery	Light gray	Yellow	−	+	+
EAM 15	Good	Powdery	Light gray	Pale yellow	−	+	+
EAM 16	Good	Leathery	Light brown	Brown	Pale brown	+	+
EAM 17	Good	Powdery	White	Yellow	Yellow	+	+
EAM 18	Good	Powdery	Gray	Brown	Brown	+	+
EAM 19	Good	Powdery	Gray	Brown	Light yellow	+	+
EAM 20	Good	Powdery	Gray	Brown	−	+	+
EAM 21	Good	Powdery	Gray	Brown	Pale yellow	+	+
EAM 22	Good	Powdery	Light gray	Yellow	Light gray	+	+
EAM 23	Good	Powdery	Gray	Pale yellow	Pale yellow	+	+
EAM 24	Good	Powdery	Gray	Yellow	−	+	+
EAM 25	Good	Powdery	Gray	Brown	−	+	+
EAM 26	Good	Powdery	Brown	Yellow	Brown	+	+
EAM 27	Good	Powdery	White	Brown	Gray	+	+
EAM 28	Good	Powdery	Dark brown	Grey	Brown	+	+
EAM 29	Good	Powdery	Gray	Brown	Gray	+	+
EAM 30	Good	Powdery	Gray	Brown	Gray	+	+

+: Present; −: Absent, AMC: Aerial Mass Color, RSP: Reverse Side Pigment, SP: Soluble Pigment, AM: Aerial Mycelium, SM: Substrate Mycelium

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

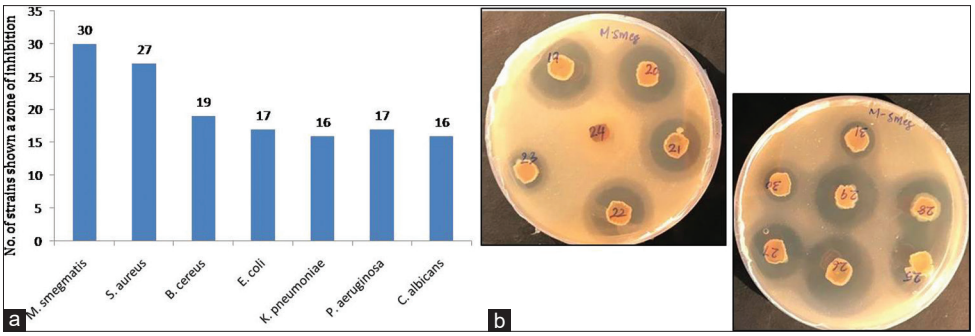


Figure 1: Antimicrobial activity of actinobacterial strains. (a). No. of actinobacteria strains showing the inhibition against various pathogens. (b). Zone of inhibition against *Mycobacterium smegmatis*

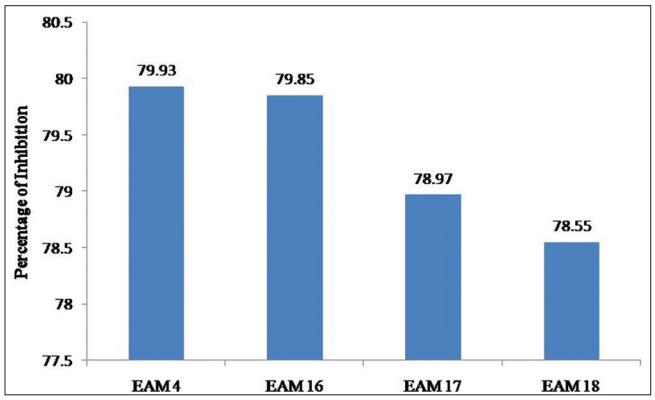


Figure 2: Antimycobacterial activity of endophytic actinobacteria

were found to be active against Gram-negative bacteria with or without showing activity against Gram-positive bacterial pathogens. It also noted that in a previous study by Manikkam *et al*,[17] 50% of the actinobacterial cultures isolated from rare ecosystems in India showed activity against Gram-negative bacterial pathogens.

Screening for antimycobacterial activity using *Mycobacterium smegmatis*

In this study, *M. smegmatis* culture was inhibited by 28 actinobacterial cultures. The zone size was ranged between 10 and 25 mm in diameter [Figure 1]. Notably, 16 actinobacterial strains showed zone of inhibition between 20 and 25 mm in diameter [Figure 2]. It is difficult to use slow-growing mycobacterium like MTB for screening a huge number of compounds or extracts on a high throughput basis. Therefore, in previous studies, *M. smegmatis* is the surrogate of choice for antimycobacterial screening due to its nonpathogenic nature, faster growth rate than MTB. The bioassay using *M. smegmatis* as a surrogate usually serves to prioritize the candidates in anti-TB screening.[28] The active extracts or compounds further can be tested in more specific *in vitro* assays on pathogenic drug-sensitive, MDR-, and XDR-MTB strains. In some previous studies, extracts of endophytic actinobacterial strains are reported to be active against *M. smegmatis*,[29] whereas in the present study, the endophytic actinobacterial cultures were screened by the simple agar plug method in which 93% of actinobacterial strains were found to be active against *M. smegmatis*. Hence, this agar plug method is a simple and rapid approach to study the antimycobacterial activity of Actinobacteria using *M. smegmatis* as a substitute host.

Production, extraction, and activity testing

In agar surface fermentation, the actinobacterial culture EAM4 showed good growth on YEME agar medium. Among the four different solvents tested, crude extracts were obtained only in the ethyl acetate (21 mg) and methanol (29 mg) but not in the chloroform and n-hexane. In the agar well diffusion method, only the ethyl acetate extract showed inhibition against *S. aureus* (16 mm), *E. coli* (10 mm), and *M. smegmatis* (23 mm). The activity exhibited by the methanol extract was not considered since it showed less than a 10 mm zone of inhibition. The previous study also reported that the ethyl acetate extract of actinobacterial strains was isolated from various sources which showed the maximum level of antimycobacterial activity.[29,30]

In vitro screening of strain EAM4 ethyl acetate extracts against *M. tuberculosis*

In the antitubercular screening by the LRP assay, the ethyl acetate extract of strain EAM4 showed 79.93% inhibition against the standard laboratory strain MTB H37Rv. The present investigation is exactly correlated with other studies, showing that ethyl acetate extract of endophytic actinobacteria isolated from medicinal plant showing anti-TB activity against MTB H37Rv.[29,31]

Characterization and identification of strain EAM4 by 16S rRNA analysis

Strain EAM4 produced good colonial growth with powdery consistency with brown soluble and reverse side pigment on the YEME agar medium. Both the aerial and substrate mycelium were observed under the microscopic observation. PCR amplification of the 16S rRNA gene from the strain EAM4 yielded 1491 base pair sequence size. The BLAST analysis revealed that the EAM4 16S rRNA gene sequence showed 99.87% similarity with the genus *Streptomyces*. Phylogenetic tree constructed based on the neighbor-joining method also indicated its closest similarity with *Streptomyces maritimus* and *Streptomyces rochei* (99%) and other *Streptomyces* species [Figure 3]. The 16S rRNA gene sequence of strain

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

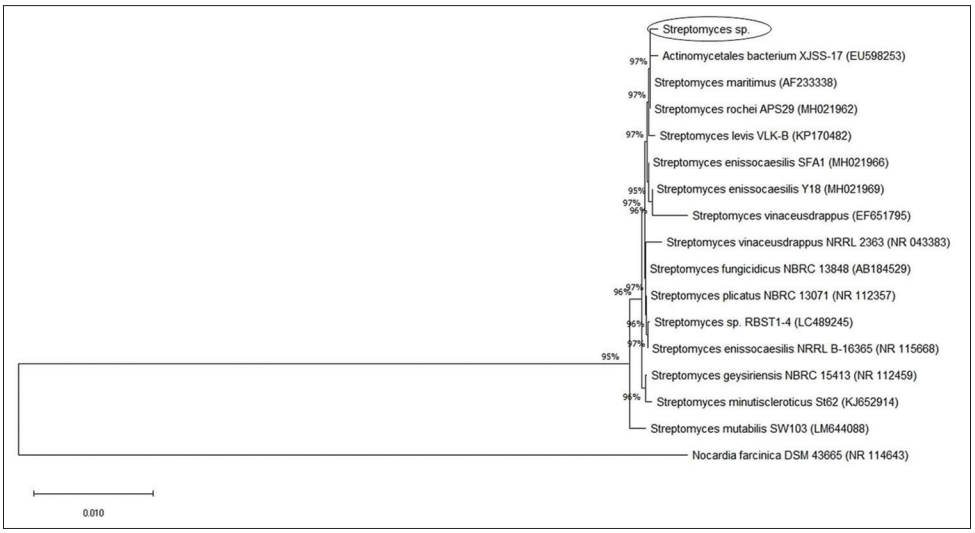


Figure 3: Neighborjoining phylogenetic tree based on 16SrRNA gene sequences comparing strain EAM4 to *Streptomyces* species. The number on the branches indicates the percentage bootstrap values of 1000 replicates; only values >50% are indicated. *Nocardia farcinica* DSM43665 used as outgroup

Table 3: Antimicrobial activity of endophytic actinobacteria by agar plug method

Strains	*Staphylococcus aureus*	*Bacillus cereus*	*Escherichia coli*	*Klebsiella pneumoniae*	*Pseudomonas aeruginosa*	*Candida albicans*
EAM 1	12	0	0	0	12	0
EAM 2	20	17	15	12	0	20
EAM 3	19	15	15	12	12	20
EAM 4	15	11	10	10	10	10
EAM 5	13	0	0	0	12	0
EAM 6	15	16	13	10	18	20
EAM 7	11	9	0	0	0	0
EAM 8	10	0	0	0	10	0
EAM 9	16	12	14	10	0	12
EAM 10	10	0	0	0	0	0
EAM 11	21	18	10	15	0	20
EAM 12	16	12	14	10	18	12
EAM 13	0	0	0	0	0	0
EAM 14	14	0	0	0	12	0
EAM 15	0	0	0	0	0	0
EAM 16	13	12	0	0	12	0
EAM 17	16	11	14	10	11	12
EAM 18	18	15	10	12	0	10
EAM 19	18	0	11	12	0	12
EAM 20	0	17	0	0	0	0
EAM 21	19	0	15	15	0	21
EAM 22	10	11	0	0	12	0
EAM 23	15	16	0	0	0	10
EAM 24	19	16	11	15	0	18
EAM 25	19	0	11	13		23
EAM 26	10	0	21	0	0	0
EAM 27	19	15	15	12		20
EAM 28	12	10	0	0	0	0
EAM 29	20	17	15	12	0	20
EAM 30	15	14	0	0	0	0

EMA4 was published in GenBank with the accession number KJ413010. There are several antitubercular antibiotics that are reported from different *Streptomyces* sp. but not from *S. maritimus* and *S. rochei*. However, in the present study,

[Downloaded free from http://www.bmbtrj.org on Tuesday, January 31, 2023, IP: 14.139.184.28]

antitubercular activity was exhibited by the *Streptomyces* EAM4 which is close to *S. maritimus* and *S. rochei.* Moreover, the production of secondary metabolites is a strain-specific process rather than a genus and species specific.

Conclusion

Findings of the present study revealed that the endophytic cultures of *Streptomyces* investigated in this study are the promising source for isolating antimycobacterial metabolites effective against tuberculous and NTM pathogens.

Acknowledgement

The authors would like to thank the authorities of Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, for the research facilities provided.

Financial support and sponsorship

This work was supported by the management of Sathyabama Institute of Science and Technology, Chennai, by providing the research facilities.

Conflicts of interest

There are no conflicts of interest.

References

1. World Health Organization. Global Tuberculosis Report 2019. Available from: https://www.who.int/tb/publications/global_report/en/. [Last accessed on 2020 Jan 07].
2. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, *et al*. Tuberculosis. Nat Rev Dis Primers 2016;2:16076.
3. Yuan T, Sampson NS. Hit generation in TB drug discovery: From genome to granuloma. Chem Rev 2018;118:1887-916.
4. Ratnatunga CN, Lutzky VP, Kupz A, Doolan DL, Reid DW, Field M, *et al*. The rise of non-tuberculosis mycobacterial lung disease. Front Immunol 2020;11:303.
5. Johansen MD, Herrmann J, Kremer L. Non-tuberculous mycobacteria and the rise of *Mycobacterium abscessus*. Nat Rev Microbiol 2020;18:392-407.
6. Crowley GC, O'Mahony J, Coffey A, Sayers R, Cotter P. A rapid viability and drug-susceptibility assay utilizing mycobacteriophage as an indicator of drug susceptibilities of Anti-TB drugs against *Mycobacterium smegmatis* mc 155. Int J Mycobacteriol 2019;8:124-31.
7. Radhakrishnan M, Jerrine J, Anbarasu S, Suresh A, Balagurunathan R. Actinobacteria from less explored ecosystems: A promising source for anti TB metabolites. (Book Chapter). Curr Res Microbiol 2017:1-29.
8. Quan D, Nagalingam G, Payne R, Triccas JA. New tuberculosis drug leads from naturally occurring compounds. Int J Infect Dis 2017;56:212-20.
9. Nguta JM, Appiah-Opong R, Nyarko AK, Yeboah-Manu D, Addo PG. Current perspectives in drug discovery against tuberculosis from natural products. Int J Mycobacteriol 2015;4:165-83.
10. Igarashi M. Development of new antituberculosis drugs from natural products. Biosci Biotechnol Biochem 2017;81:32-7.
11. Grzelak EM, Choules MP, Gao W, Cai G, Wan B, Wang Y, *et al*. Strategies in anti-*Mycobacterium tuberculosis* drug discovery based on phenotypic screening. J Antibiot (Tokyo) 2019;72:719-28.
12. Golinska P, Wypij M, Agarkar G, Rathod D, Dahm H, Rai M. Endophytic actinobacteria of medicinal plants: Diversity and bioactivity. Antonie Van Leeuwenhoek 2015;108:267-89.
13. Sharma P, Thakur D. Antimicrobial biosynthetic potential and diversity of culturable soil actinobacteria from forest ecosystems of Northeast India. Sci Rep 2020;10:4104.
14. Christina A, Christapher V, Bhore SJ. Endophytic bacteria as a source of novel antibiotics: An overview. Pharmacogn Rev 2013;7:11-6.
15. Strobel G, Daisy B. Bioprospecting for microbial endophytes and their natural products. Microbiol Mol Biol Rev 2003;67:491-502.
16. Vijitha V, Radhakrishnan M, Balagurunathan R. Bioprospecting of endophytic actinomycetes for antiphytofungal activity. Int J ChemTech Res 2014;6:2689-694.
17. Manikkam R, Venugopal G, Subramaniam B, Ramasamy B, Kumar V. Bioactive potential of actinomycetes from less explored ecosystems against *Mycobacterium tuberculosis* and other nonmycobacterial pathogens. Int Sch Res Notices 2014;2014. doi: 10.1155/2014/812974.
18. Pazhanimurugan R, Radhakrishnan M, Shanmugasundaram T, Gopikrishnan V, Balagurunathan R. Terpenoid bioactive compound from Streptomyces rochei (M32): Taxonomy, fermentation and biological activities. World J Microbiol Biotechnol 2016;32:161.
19. Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406-25.
20. Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985;39:783-91.
21. Ahmed N, Ayoub K, Asma A, Yedir O, Lahcen H. Endophytic actinobacteria of medicinal plant Aloe vera: Isolation, antimicrobial, antioxidant, cytotoxicity assays and taxonomic study. Asian Pac J Trop Biomed 2018;8:513-18.
22. Shenpagam HN, Kanchana D, Sinduja G, Sandhya R. Isolation of endophytic actinomycetes from medicinal plants and its mutational effect in biocontrol activity. Int J Pharmacet Sci Res 2012;3:4338-344.
23. Madhurama G, Sonam D, Urmil PG, Ravindra NK. Diversity and biopotential of endophytic actinomycetes from three medicinal plants in India. African J Microbiol Res 2014;8:184-91.
24. Anusree T, Bhai RS. Rhizosphere actinobacteria for combating Phytophthora capsici and Sclerotium rolfsii, the major soil borne pathogens of black pepper (Piper nigrum L.). Biol Control 2017;109:1-13.
25. Salam N, Khieu TN, Liu MJ, Vu TT, Chu-Ky S, Quach NT, *et al*. Endophytic actinobacteria associated with Dracaena cochinchinensis Lour: Isolation, diversity, and their cytotoxic activities. Biomed Res Int 2017;2017. doi: 10.1155/2017/1308563.
26. Sharma P, Baunthiyal M. Endophytic Actinobacteria from Pinus roxburghii: Isolation, diversity and antimicrobial potential against human pathogens. J Pharmacognosy Phytochem 2018;7:3021-7.
27. Bérdy J. Thoughts and facts about antibiotics: Where we are now and where we are heading. J Antibiot (Tokyo) 2012;65:385-95.
28. Zamakhaev M, Tsyganov I, Nesterova L, Akhova A, Grigorov A, Bespyatykh J, *et al*. *Mycolicibacterium smegmatis* possesses operational agmatinase but contains no detectable polyamines. Int J Mycobacteriol 2020;9:138-43.
29. Khunjamayum R, Tamreihao K, Debananda SN, Athokpam S, Salam N, Sharmistha B, *et al*. *In-vitro* antimycobacterial activities of endophytic bacteria associated with medicinal plant of Manipur. J Bacteriol Mycol 2017;4:104-107.
30. Manigundan K, Revathy S, Sivarajan, Anbarasu S, Jerrine J, Radhakrishnan M, *et al*. Bioactive potential of selected actinobacterial strains from Andaman marine ecosystems effective against *Mycobacterium tuberculosis* and other clinical pathogens. Indian J Geo Marine Sci 2019;48:1307-1311.
31. Radhakrishnan M, Suganya S, Balagurunathan R, Vanaja K. Preliminary screening for antibacterial and antimycobacterial activity of actinomycetes from less explored ecosystems. World J Microbiol Biotechnol 2010;26:561-66.

Antimicrobial and anti-tubercular potential of Streptomyces lomondensis SACC 63 isolated from soil samples in Himachal Pradesh, India

Soundarya, G.; K., Manigundan; P.R., Meganathan; S., Gomathi; Balaji, S.; Jerrine, J.; M., Radhakrishnan.

J Environ Biol; 2020 Jan; 41(1): 131-138
Article | IMSEAR | ID: sea-214484

ABSTRACT

Aim:

The antimicrobial and anti-tubercular potential of Actinobacteria isolated from soil in Himachal Pradesh, India, was examined in this study.

Methodology:

The bioactive metabolites produced from five selected Actinobacterial cultures SACC-63, SACC-75, SACC-76, SACC-77 and SACC-N10 by agar surface fermentation were tested for their antimicrobial and anti-tubercular efficacy in in-vitro. The taxonomy of SACC-63 was identified as Streptomyces lomondensis using 16S-rRNA sequence analyses.

Results:

The ethyl acetate extract of strain SACC-63 showed broad spectrum activity against wide range of bacterial and fungal pathogens at 250 µg per disc concentration. In anti-tubercular screening by luciferase reporter phage (LRP) assay, four actinobacterial strains showed more than 65% inhibition against the standard strain Mycobacterium tuberculosis H37Rv and multi drug resistant (MDR) M. tuberculosis strain at 100 µg ml-1 and 500 µg ml-1 concentrations, except the strain SACC-N10. The MIC values of SACC-63 against microbial pathogens ranged between 62.5 µg ml-1 and 250 µg ml-1. The strain SACC-63 showed 99% similarity with Streptomyces spp. Further, phylogenetic analysis based on 16S rRNA revealed a close relationship between SACC-63 and S. lomondensis.

Interpretation:

The present study suggested that S. lomondensis SACC-63 isolated from Himachal Pradesh region in India could be a promising source for the production of bioactive molecules.

Antimicrobial; Anti-tuberculosis; Bioprospecting; Luciferas

screening

Organized periodic procedures performed on large groups of people for the purpose of detecting disease.

Source DeCS - Health Sciences Descriptors

WE WISH YOUR OPINION

Microbiological Communication

BBRC Bioscience Biotechnology Research Communications

Biosc.Biotech.Res.Comm. Vol 13 (3) July-Aug-Sep 2020 Pp-1050-1054

Antimicrobial and Antitubercular Activity of Endophytic Actinobacterium, *Streptomyces* sp. SACC 4 Isolated from the Mangrove Plant *Rhizophora apiculata*

Ayswarya, S., K. Manigundan, G. Vijayalashmi, A. Vignesh, V. Gopikrishnan, S. Anbarasu, Jerrine Joseph, and M. Radhakrishnan*

Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, Tamil Nadu – 600 119. India

ABSTRACT

Public health threat due to drug resistant pathogens like *Mycobacterium tuberculosis* and others necessitates the discovery of novel antibiotics. Actinobacteria from understudied sources is emerging as promising source for novel molecules effective against drug resistant pathogens. The present study was undertaken for bioprospecting of endophytic actinobacteria from the mangrove plant *Rhizophora apiculata*. Different actinobacterial cultures isolated from *Rhizophora apiculata* leaves were screened for antimicrobial activity by agar plug method against a wide range of bacterial and fungal pathogens. Antimicrobial metabolites from potential strain SACC4 was produced by agar surface fermentation and extracted using different organic solvents and tested against bacterial pathogens. Antitubercular activity of strain SACC4 was also tested by adopting luciferase reporter phage (LRP) assay. Strain SACC4 was identified based on their phenotypic and molecular characteristics. About 28 distinct actinobacterial strains were isolated from the mangrove leaves in which 50% of the strains showed antimicrobial activity. Notably strain SACC 4 showed broad spectrum activity against bacterial pathogens. Based on the studied phenotypic and molecular characteristics, strain SACC4 was identified as *Streptomyces* species. Among the different solvents tested, methanol extract showed a maximum antimicrobial activity against *S. aureus* in disc diffusion method at 100µg/ml concentration. The methanol extract showed more than 60 percentage inhibition against the standard strain *M. tuberculosis* H37Rv, drug sensitive and MDR *M. tuberculosis* strains at both 100µg/ml and 500µg/ml concentrations. The results of the present investigation revealed that the endophytic *Streptomyces* from mangrove plants are the potent source of novel bioactive metabolites effective against *M. tuberculosis* and other drug resistant pathogens.

KEY WORDS: ACTINOBACTERIA, STREPTOMYCES, ANTIMICROBIAL, ANTI TB, TAXONOMY.

Article Information:
Corresponding author email: *mrkactinos@gmail.com*
Received 15/07/2020 Accepted after revision 10/09/2020
P-ISSN: 0974-6455 E-ISSN: 2321-4007
Thomson Reuters ISI Clarivate Analytics
Web of Science ESCI Indexed Journal

Identifiers & Pagination:
Vol 13(3) E-Pub 30th Sep 2020 Pp- 1050-1054
This is an open access article under Creative Commons License Attribution International (CC-BY 4.0)
Published by Society for Science & Nature India
DOI: *http://dx.doi.org/10.21786/bbrc/13.3/9*

INTRODUCTION

Antimicrobial resistance (AMR) is a global health crisis. Multi drug resistant pathogens like *Mycobacterium tuberculosis* pose serious threat to public health. Hence, the critical component of AMR solution is the development of truly novel antibiotic. Natural products and their derivatives are mainstays of our antibiotic drugs (Wright, 2017) in which members of the bacterial phylum actinobacteria in particular the genus *Streptomyces* are the known produces of numerous novel secondary metabolites especially antibiotics. However, in recent years, searching of routine sources like terrestrial soil results in the isolation of known actinobacteria which in turn produce known bioactive metabolites (Berdy, 2012). Instead, exploring rare sources like plants may results in the isolation of novel actinobacterial strains with promising bioactive potential. Actinobacteria reside within the healthy plants as endophytes are the promising source for bioactive metabolites to fight against drug resistant pathogens (Golinska et al., 2015, Simpkin et al., 2017). Limited reports on endophytic actinobacetria from plants revealed their enormous bioprospecting potential for clinical, environmental and agricultural applications, (Golinska et al., 2015; Singh et al., 2018).

The present study reports the isolation of endophytic actinobacteria from the mangrove plant *Rhizophora apiculata* and their antimicrobial and antitubercular properties.

MATERIAL AND METHODS

Leaves sample of the mangrove plant *Rhizophora apiculata* was collected from Parangipettai mangrove zone (11.50°N, 79.75°E), Tamil Nadu, India. Endophytic actinobacteria from the leaves were isolated by adopting the method described by Golinska et al., (2015). Morphologically distinct actinobacterial strains were recovered and preserved on ISP2 agar slants at 4°C. All the actinobacterial cultures were screened for antimicrobial activity by agar plug method against clinical strains of *Staphylococcus aureus* MTCC96, clinical strains of *Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa*, Candida albicans and *Aspergillus fumigates* (Radhakrishnan et al., 2014). Actinobacterial strain SACC 4 which showed promising activity was selected as potential strain.

Phenotypic and molecular characteristics of strain SACC4 was studied by the methods described by Shirling and Gottileb (1966) and Radhakrishnan et al., (2013). The 16S rRNA gene of SACC4 was amplified by using the primer pairs: 27F 5′AGAGTTTGATCMTGGCTCAG3′ (forward) and 1492R 5′TACGGYTACCTTGTTACGACTT3′ (reverse) and sequencing was carried out at Eurofins Genomics India Pvt. Ltd., Bangalore. The identification of phylogenetic neighbors and calculation of pair wise 16SrRNA gene sequence similarities were achieved using the MEGA version 6 (Tamura et al., 2013) and BLAST analysis (http://blast.ncbi.nlm. nihgov/Blast.cgi). The obtained 16SrRNA sequence was submitted to GenBank to get the accession number.

Effect of solid-state and submerged fermentation on bioactive metabolite production by the strain SACC4 was investigated against *S. aureus* by agar well diffusion method. Effect of solvents on the extraction of bioactive metabolites from the cell free supernatant of strain SACC4 was studied using different solvents such as n-hexane, chloroform and ethyl acetate at 1:2 ratio for 24 hours. The antibacterial activity of different solvent extracts was tested against *S. aureus* MTCC96 by disc diffusion method (Bavya et al., 2011). The antimicrobial activity of methanol extract of strain SACC4 was evaluated against wide range of bacterial pathogens by disc diffusion method at 100 µg/disc concentration. Zone of inhibition was measured after 24 hours of incubation at 37°C and expressed in millimetre in diameter. Effect of different concentration of YEME medium (2X, 1X, 1/2X, 1/4X and 1/10X) on antimicrobial activity of strain SACC 4 was also tested (Yilmaz et al., 2008).

Antitubercular activity of potential strain of SACC4 was studied against standard laboratory strain *Mycobacterium tuberculosis* H37Rv, drug sensitive and multi drug resistant (MDR) clinical isolates of *M. tuberculosis* by adopting LRP assay. The relative light unit (RLU) was measured in a luminometer (Radhakrishnan et al., 2014).

Percentage RLU reduction = Control RLU – Test RLU / Control RLU X 100

Extract showing RLU reduction by 50% or more when compared to control were considered as having anti-tubercular activity.

RESULTS AND DISCUSSION

Totally 28 endophytic actinobacterial cultures were isolated from the *Rhizophora apiculata* leave. Based on their morphology, about 75 % of the cultures were tentatively identified as *Streptomyces* and the remaining 25% are rare actinobacterial strains. Actinobacteria from understudied ecosystems are the promising source for novel bioactive natural products. Limited reports on endophytic actinobacetria from plants revealed their enormous bioprospecting potential for clinical, environmental and agricultural applications (Golinska et al., 2015; Singh et al., 2018). In India, reports on endophytic actinobacteria are very few. In a previous study, Du et al. (2013) analyzed the reported 600 actinobacteria belonging to 34 genera and 7 unknown taxa from 37 medicinal plants in which maximum from root followed by stem and least from leaves. However in the present study 25% of the actinobacterial strains recovered from leaves were found to be rare actinobacterial strains.

In preliminary screening, 18 actinobacterial cultures (64%) showed inhibition against any one of the bacterial pathogens tested. Notably, strain SACC 4

showed promising inhibition against *S. aureus* with 15-18 mm zone of inhibition against *S. aureus, E. coli*, *K. pneumoniae, P. aeruginosa* and the fungi *C. neoformans*.

Microscopic, cultural and physiological characteristics of strain SACC4 was given in table 1. PCR amplification and BLAST analysis revealed that the SACC 4 sequence showed 99.9% similarity with the genus *Streptomyces*. Phylogenetic tree constructed based on neighbor-joining method also indicated its closest similarity with *Streptomyces djakartensis* (99%) and other species (Figure 1). The 16s rRNA gene sequence of strain SACC 4 was published in GenBank with the accession number KJ413010. Strain SACC4 was showed 99.8% close similarity with the 16SrRNA sequence of Streptomyces djakartensis. However, certain differences were noted when comparing the physiological characteristics of strain SACC 4 with the characteristics of S. djakartensis (Kounouz et al., 2017). Hence it was assumed that the strain SACC 4 is found to be novel at strain level. In another previous study, it was reported that 85 % of the total 123 endophytic actinobacterial isolates studied were determined to be unique at the strain level (Janso et al., 2010). Further DNA-DNA hybridization and G+C content analysis is needed to confirm its novelty.

Table 1: Characteristics of endophytic actinobacterial strain SACC 4

	Characteristics	SACC4
Micromorphology	Aerial mycelium	+
	Substrate mycelium	+
Cultural characteristics	Fragmentation	-
	Colony consistency	Powdery
	Aerial mass colour	Gray
	Reverse side pigment	-
	Soluble pigment	-
Physiological Characteristics	ISP1 (Tryptone Yeast Extract Agar)	-
	ISP2 (Yeast extract malt extract agar)	+
	ISP3 (Oat meal agar)	-
	ISP4 (Inorganic salt starch agar)	+
	ISP5 (Glycerol asparagine agar)	-
	ISP6 (Peptone yeast Iron agar)	+
	ISP7 (Tyrosine agar)	+
Carbon utilization	Arabinose	+
	Xylose	+
	Inositol	+
	Mannitol	+
	Fructose	+
	Rhamnose	+
Enzyme production	Asparagine	+
	Glutamine	+
PH tolerance	5	-
	6	-
	7	+
	8	+
	9	+
	10	+
	11	+
Temperature tolerance (° C)	20	+
	30	+
	40	+
NaCl tolerance (%)	0	+
	1	+
	2	+
	3	+
	5	-

"+" grown, " – " no growth

Figure 1.The phylogenetic relationship of the potential actinobacterial strain SACC4 based on 16S rRNA gene homology. The tree was constructed using the neighbor-joining method with pairwise-deletion model analyses, which were implemented in the Molecular Evolutionary Genetics Analysis (MEGA), version 6.0 program. The resultant tree topologies were evaluated by bootstrap analysis based on 1000 replicates. Micromonosporainyonens is was used as out group. Scale bar indicates the number of substitutions per site.



In fermentation study, even until 20 days of fermentation, strain SACC 4 was failed to produce bioactive compounds in submerged fermentation whereas it produced bioactive compounds in agar surface fermentation. The agar plug from ISP2 agar showed activity against *S. aureus* even from 2nd day (11.2±0.72 mm in dm) of fermentation. The zone of inhibition eventually increased from 2nd day to 8th day (17.1±1.09 mm in dm) of fermentation. This is due to the antibiotics production through solid state fermentation are higher quantities and more stable than liquid state fermentation (Subramanian et al., 2012). This finding is in agreement with the antibiotic production by desert soil, Streptomyces strain D25 which showed antibacterial activity and also produced antitubercular pigment only in solid culture (Radhakrishnan et al., 2015).

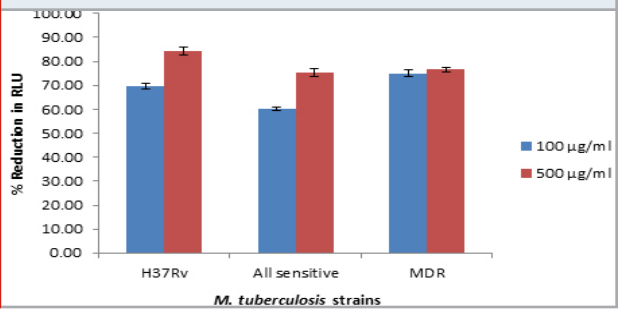
In solid state fermentation, strain SACC4 showed good growth and the methanol extract obtained from the fermented medium showed 18mm inhibition against *S. aureus* but failed to show activity in liquid state fermentation. Isolation of crude extract by solvent extraction is an important phenomenon to find a good solvent that has the potential to extract high yield and most potent bioactive compounds. Our research findings demonstrated that the extract of methanol have significant antimicrobial and antitubercular activity. These results were clearly comparable to that, who evaluated Streptomyces mutabilis which produced an antitubercular activity and antimicrobial activity against gram positive and gram negative pathogens (Mahmoud et al., 2015). Among the different concentrations of YEME agar, 1X concentration of YEME agar was found to support good growth and antimicrobial compound production which was expressed by its activity against *S. aureus* (10.2±0.12), *B. subtilis* (15.0±0.64) and *E. coli* (9.8±0.44). The methanol extract of strain SACC 4 was found to be active against *S. aureus* and wide range of Gram negative bacterial pathogens (Table 2).

Table 2. Antimicrobial activity of methanol extract of endophytic actinobacterial strain SACC 4.

Bacterial pathogens	Zone of inhibition (mm in diameter)*
Staphylococcus aureus	18.2 ±0.15
Klebsiella pneumonia	17.0
Pseudomonas aeruginosa	18.0
Vibrio cholerae 01 ogawa	8.1±0.17
Vibrio cholerae 0139	10.2±0.15
Vibrio parahaemolyticus 03:K6	12.8±0.47
Aeromonassorbia	8.3±0.17
Aeromonashydrophilla	10.2±0.15
Enteropathogenic E.coli 0115	10.2±0.37
Enteroaggregative E.coli 015	10.9±0.35
Enteropathogenic E.coli 0114	10.9±0.21
Salmonella typhimurium	12.0±0.32
S. worthington	12.9±0.38
S. infantis	11.4±0.31
Shigella sonnei	0
S. dysenteriae sero 5	0
S.. boydii sero 1	0
S. flexneri type 2a	0

In LRP assay, the methanol extract of strain SACC 4 was showed more than 60% inhibition against standard laboratory strain *Mycobacterium tuberculosis* H37Rv, drug sensitive and MDR *M. tuberculosis* strains (Figure 2).

Figure 2: Antitubercular activity of potential actinobacterial strainSACC4



Findings of the present study evidenced that mangroves are the promising source for the isolation of bioactive endophytic actinobacteria.

ACKNOWLEDGEMENTS

We thank the management of Sathyabama Institute of Science and Technology (Deemed to be University), Chennai for the research facilities provided. Authors (Dr MRK and AS) thank the DST-SERB for their support in

the form of research grant (Ref:YSS/2015/001887) under Young Scientist Scheme.

Conflict of Interest: Authors have no conflict of interest.

REFERENCES

Bavya M, Mohanapriya P, Pazhanimurugan R, Balagurunathan R.(2011) Potential bioactive compound from marine actinomycetes against biofouling bacteria. Indian J Geo-Marine Sci 40(4): 578-582.

Berdy J.(2012) Thoughts and facts about antibiotics: where we are now and where we are heading. J Antibiot 65(8): 385-395.

Du H, Su J, Yu L, Zhang Y.(2013) Isolation and physiological characteristics of endophytic actinobacteria from medicinal plants. Wei Sheng Wu Xue Bao 53(1):15-23.

Golinska P, Wypij M, Agarkar G, Rathod D, Dahm H, Rai M (2015) Endophytic actinobacteria of medicinal plants: diversity and bioactivity. Antonie van Leeuwenhoek 108: 267-289.

Janso JE, Carter GT.(2010) Biosynthetic potential of phylogenetically unique endophytic actinomycetes from tropical plants. Appl Environ Microbiol 76: 4377-4386.

Kounouz Rachedi, Ferial Zermane, Radja Tir, Fatima Zohra Ayache, Robert Duran, Béatrice Lauga, Solange Karama , Maryse Simon and Boulahrouf Abderrahmane (2017). Characterization of tribenuron-methyl resistant Actinobacteria strains. Asian Journal of Microbiology, Biotechnology and Environmental Sciences. 19. 81-87.

Mahmoud AY, Abdallah HM, El-Halawani MA, Jiman-Fatani AAM (2015) Anti-tuberculous activity of Treponemycin produced by a *Streptomyces* strain MS-6-6 isolated from Saudi Arabia. Molecules 20: 2576-2590.

Radhakrishnan M, Gopikrishnan V, Balaji S, Balagurunathan R, Vanaja K.(2014) Bioactive Potential of Actinomycetes from Less Explored Ecosystems against Mycobacterium tuberculosis and Other Non mycobacterial Pathogens. Int Scholarly Res Notice Article ID 812974.

Radhakrishnan M, Gopikrishnan V, Suresh A,Selvakumar N, Balagurunathan R, Vanaja Kumar (2013) Characterization and phylogenetic analysis of antituberculous compound producing actinomycete strain D25 isolated from Thar Desert soil, Rajasthan. Bioinformation 9: 018-022.

Radhakrishnan M, Venugopal G, Ramasamy B, Kumar V.(2015) Effect of critical medium components and culture conditions on anti-tubercular pigment production from novel Streptomyces sp D25 isolated from Thar dessert, Rajasthan. J App Pharm Sci 5(06): 015-019.

Shirling BE, Gottlieb D.(1966) Methods for characterizing Streptomyces species. Int J Syst Bacteriol 16: 313-340.

Simpkin VL Matthew JR, Ruth K, Elias M.(2017) Incentivising innovation in antibiotic drug discovery and development: progress, challenges and next steps. J Antibiot 70: 1087-1096.

Singh R, Dubey AK.(2018) Diversity and applications of endophytic actinobacteria of plants in special and other ecological niches. Front Microbiol 9; 1767.

Subramaniyam R, Vimala R.(2012) Solid state and submerged fermentation for the production of bioactive substances: a comparative study. Int J Sci Nat 3: 480-486.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30: 2725-2729.

Wright GD (2017) Opportunities for natural products in 21st century antibiotic discovery. Nat Prod Rep 34: 694-701.

Yilmaz EB, Yavuz M, Kizil M. (2008) Molecular characterization of rhizosphere soil *Streptomycetes* isolated from indigenous Turkish plants and their antimicrobial activity. World J Microbiol Biotechnol 24: 1461-1470

Indian Journal of Environmental Protection

December-2020

IJEP 40 (12) : 1247-1254 (2020)

Study Of Ambient Air Quality On Transport Network Near Mining Region- A Case Of Talcher Coalfield

Krishna Prakash, Vikas Nimesh, Md. Saddam Hussain and Arkopal Kishore Goswami

Abstract

Dust generation due to movement of heavy trucks on transport and haul roads is one of the major sources of air pollution in open cast coal mining areas. Dust particles generated by mining operations degrade air quality which causes serious health problems. Talcher is heavily polluted due to heavy truck movement, industrial activity and extensive mining operations. Significant proportions of air pollutants are released from industries and vehicles, among which respirable particulate matter (PM10) is a major concern. This research aims to understand the impact of movement of trucks on the overall air quality in the vicinity of a coal mine. The correlation between air pollutants and meteorological factors in these areas has also been analyzed. Study area was selected on the basis of type of mining, accessibility, method of working, meteorological condition and resource availability. The concentration of particulate matter (PM10) was monitored in specified location for 5 days (minimum of 4 hr daily) by using EVM 7. Micro-meteorological parameters, like air temperature and humidity, were collected by using WatchDog 2900ET weather station and EVM 7. The results were then compared with National Ambient Air Quality Standards (NAAQS). It was observed that during the study period, the concentration levels of PM10 generally exceeded the permissible standard. The concentration of PM10 ranged from 20-782 g/m3 with an average of 211 g/m3. The influence of relative humidity and temperature on the concentration of PM10, CO and CO2 were analyzed. The result of the analysis showed that PM10 had very weak and negative correlation with temperature during the monitoring period. CO had strong positive relationship with temperature and significant negative correlation with relative humidity during study period. The results also showed that the correlation between PM10 and relative humidity was very weak (negative) during monitoring period (post-monsoon season).

Keywords

Air quality, PM10, Temperature, Humidity, Transport road, Opencast coal mining

Privacy - Terms

View Full Length Paper

IJEP 40 (12) : 1255-1263 (2020)

Studies On The Geomorphology And Land Use / Land Cover (LU/LC) Analysis On Annamalai Hill, Tiruvannamalai District, Tamil Nadu

M. Pandian, S. Natarajan and V. Madha Suresh

Abstract

Geomorphology and landuse/land cover (LU/LC) changes of Annamalai hill and surrounding urban areas, Tiruvannamalai district, Tamil Nadu, were determined from 2006-2018 using remote sensing technology and geographic information system (GIS). On the basis of field study, geomorphology and remote sensing data, the study area was classified into six categories. Analysis of toposheet and satellite imaging of LU/LC in 2006 and 2018 revealed that there is a significant increase in built-up area and barren wastelands during the study period due to rapid urbanization and mass tourism. The extent of vegetation cover, arable land, water bodies and hills area have substantially declined due to various anthropogenic factors, such as grazing, man-made forest fire, encroachment of forest land, increasing population, generation of garbage and mass tourism.

Keywords

Anthropogenic factors, Biodiversity, Forest fires, Tiruvannamalai hill, Vegetation cover

View Full Length Paper

IJEP 40 (12) : 1264-1272 (2020)

Noise Levels In The Modern Urban Roundabout

Jelena Bozic, Predrag Ilic and Svetlana Ilic

Abstract

Theoretical analysis and bibliographical review of the studies relating to the noise effects of traffic management devices conclude that correctly designed roundabouts can reduce noise locally from to influence on traffic noise pollution. The results of the traffic noise mapping of a roundabout are discussed based on the measured data. Comparison of the obtained results

with the degree and duration of the exposure of population to excessive noise limit values defined by the national legislation indicates that values of the traffic noise levels at the analysed roundabout in Banja Luka are unexpectedly high. Recommendations for reducing noise levels are given in the conclusions of this paper.

Keywords

Roundabout, Road, Traffic noise, Mapping, Measurements

View Full Length Paper

IJEP 40 (12) : 1273-1279 (2020)

Anticorrosive Performance Of Phyllanthus aamarus And Senna auriculata Extract-Based Hybrid Coating On Steel

Alagesan Annadurai, Elakkiah Chinnadurai, Venugopal Gopikrishnan, Jerrine Joseph and Manikkam Radhakrishnan

Abstract

Corrosion is one of the major concerns to the construction industry and low toxic corrosion inhibitors are the need of the hour. In this context, green inhibitors are obtained from plants, which are environmentally acceptable. This study focuses on the corrosion inhibition characteristics of 3 different plant extracts from Phyllanthus amarus, Senna auriculata and the hybrid of both the plants on Fe 500 grade steel immersed in acidic solutions (NaCl) using weight loss method. Phytochemical, Fourier transform infrared spectroscopy (FTIR), gas chromatography-mass spectroscopy (GC-MS) analysis and 2,2-diphenyl-1-picryl-hydrazyl hydrate (DPPH) assay were carried out to evaluate the chemical and anti-oxidant properties of plant extracts. Accelerated corrosion technique showed a higher inhibition efficiency of 9.23 with hybrid extract when compared to other samples. Phytochemical constituents showed that the extracts contain alkaloids, flavonoids, saponins and tannins. The DPPH assay, FT-IR and GC-MS analysis showed the presence of chemical compounds and strong antioxidant activity. The study concludes that the plant inhibitors are sustainable and promising alternatives to the chemical inhibitors.

Keywords

Corrosion inhibitors, Plant extracts, Accelerated corrosion test, Eco-friendly, Anti-corrosion activity

View Full Length Paper

IJEP 40 (12) : 1280-1287 (2020)

Energy And Environmental Impact Analysis Of An ANN Based Power Optimizer For Residential Solar Photovoltaic System

S.R. Revathy and V. Kirubakaran

Abstract

The increase in power production through fossil fuel contributes to the rapid rise of greenhouse gases in the atmosphere which is the major cause of global warming. Realising the abundant availability of solar resource, many countries are moving towards photovoltaic based power production which depends on the solar irradiance and temperature. Any change in irradiance and temperature affects the power output of the system, hence optimization is required to improve the efficiency of the solar photovoltaic system. This paper proposes artificial neural networks based power optimization technique to improve the overall efficiency of the system. The efficiency, tracking speed, complexity, scalability, reliability under partial shading conditions and cost of the proposed system is compared with a popular hill climbing maximum power point tracking algorithm. The proposed study has been carried out by simulating the models in MATLAB/Simulink. The proposed method increases the power output by 21% compared to incremental conductance algorithm. Based on carbon footprint analysis, the environmental impact analysis of the proposed method has been estimated and reported in this paper.

Keywords

Solar energy, Photovoltaic, Power optimizer, Artificial neural networks, Maximum power point tracking

View Full Length Paper

IJEP 40 (12) : 1288-1297 (2020)

Groundwater Suitability Assessment Using Water Quality Index And Trace Elements Distribution In Chittar Sub-basin Of Tambaraparani River, Tamil Nadu

T. Raja Kumar Jeyavel, S. Raja Vasanth Raja, P. Saravanan, B. Thirunelakandan and M. Senthilkumar

Abstract

Groundwater quality of Chittar sub-basin of Tambraparani basin has been studied using water quality index in GIS from major cations, anions data and selected trace elements. The study area lies between latitude 8050' to 9010' N and longitude 77010' to 77030' E and Hornblende Biotite gneiss, Charnockite are main rock types. The average annual rainfall is 722 mm and the depth to water table during post-monsoon varied between 1-10 m below ground level. Groundwater samples were collected from 25 open wells in the month of June 2014 and major elements of Ca2+, Mg2+, Na+, K+, HCO3-, Cl-, SO42-, NO3-, PO42- and trace elements of Zn, Pb, Cr, Fe, Cd, Cu and Mn were analysed. The study exposed that ionic concentration of cations and anions were within the permissible limit of WHO standard for drinking in most of the sample locations, nevertheless, EC, TDS, Ca2+ and Mg2+ were found above the limit in few locations. Ca+Mg; SO4+Cl hydrogeochemical facies was identified and Gibb's diagram favours the rock water interaction. Spatial maps were prepared in ArcGIS by inverse distance weighted interpolation technique. WQI was calculated by assigning weightage to physical and chemical parameters and it indicates good and moderate water quality category. Except Fe, other elements of Pb, Cr, Mn, Cu, Zn, Co, Ni and Cd were found within permissible limit of drinking water standard.

Keywords

Groundwater quality, Groundwater suitability, Water quality index, Hydrogeochemistry, Trace elements, Chittar sub-basin

View Full Length Paper

IJEP 40 (12) : 1298-1303 (2020)

Partial Pressure Of Carbon Dioxide Fluxes At Air – Sea Interface During Different Seasons In The Bay Of Bengal From New Digha To Padapur Beach, India

Anil Kumar Giri and Abhijit Das

Abstract

In the present investigation, the partial pressure of CO2 (pCO2) and carbon fluxes (FCO2) evaluate the variability and controlling factors of water quality, nutrients availability, and carbon dioxide exchange air and sea interface. The calculated partial pressure of CO2 (pCO2) and fluxes of CO2 showed decrease that overall sea behave as a sink of CO2 (-71.84±23.08

mmol/m2/d). During winter season, it was sinking of CO2, which could be attributed to higher community respiration over production as evidence by higher utilization of oxygen comparison with other seasons. However, the pCO2 was ranged between 0.70-29.05 µatm (8.08±8.15 µatm) and fluxes CO2 ranged from -94.51 to -39.26 (-71.84±23.08 mmol.C/m2/d) which was lower than the global average reported for saline ecosystem. The results showed the sea was in a better condition as compared to the other records on water quality. These results would be treated as a baseline for future extensive ecosystem health related studies and management of the sea.

Keywords

Water quality, Carbon dioxide, Interface, Nutrients, Saline ecosystem

View Full Length Paper

IJEP 40 (12) : 1304-1310 (2020)

Managing Nexus Of Development And Environmental Implications Through Sustainable Development Goals In India

Ashwani Pant and Santosh Kumar

Abstract

India and the whole world are facing challenges in mitigating environmental implications in the era of development. Being a developing country, all the three organs of government are committed to sustainable development or inclusive growth. Progressive development has led to harnessing of natural resources which in turn has led to environmental implications. Poor sectoral coordination and institutional fragmentation have triggered an unsustainable use of resources and threatened the long-term sustainability not locally but globally also, posing challenges in achieving sustainable developmental goals. The sustainable development goals (SDGs) are broad sets of common goals for fair and sustainable health from local to international community. The aim is to end poverty, protect the planet and ensure that all people enjoy peace and prosperity, now and in the future. The nexus approach through proper implementation of sustainable developmental goals can enhance understanding of various sectors and strengthen coordination among them. The paper basically will focus on: the concept of inclusive development in sustainable developmental goals and extent of inclusion; need to develop systems understanding of what supports an inclusive and sustainable development and the issues and challenges thereof.

Keywords

Environmental implications, Sustainable development, Inclusive growth, Implementation, Sustainable development goals

View Full Length Paper

IJEP 40 (12) : 1311-1316 (2020)

Biosorption Performance Of Packed Bed Column For Removal Of Chromium From Electroplating Wastewater Using Green Algae Chlorella

S. Kanchana and P. Sivaprakash

Abstract

The removal efficiency of green algae Chlorella to biosorb chromium from industrial wastewater was investigated. All experiments were conducted using fixed bed columns. Experiments were carried out as a function of flow rate (2, 4 and 6 mL/min) and bed depth (2.5, 5, 8, 10 cm) at an initial Cr (VI) concentration of 100 mg/L. The particle size of the biosorbent ranged from 150-300 µm. The breakthrough curves were obtained at different flow rates and different bed depths. The results show that the column performed well at lowest flow rate and highest bed depth. Also, column bed capacity and exhaustion time increased with increasing bed depth. The Thomas model and Yoon-Nelson model were used to analyze the experimental data and the model parameters were evaluated. Good agreement of the experimental breakthrough curves with the model predictions was observed.

Keywords

Biosorption, Chromium (VI), Chlorella, Packed bed column

View Full Length Paper

IJEP 40 (12) : 1317-1322 (2020)

Larvicidal Efficacy Of Purified Fractions Of *Blumea eriantha* (DC) Whole Plant Extracts On Larvae Of Mosquitoes

Mragendra Singh alias Mithaee Lal Gound, Deepak Meshram and Kapil K. Soni

Abstract

The present study was based on larvicidal efficacy of Blumea eriantha (DC) whole plant extracts that destroy larvae of Aedes aegypti and Anopheles stephensi mosquitoes. Mosquitoes have developed resistance against chemical pesticides. Therefore, in the present study, a new approach of control of mosquitoes was developed by applying larvicides which are biodegradable, less toxic and also facilitate the eco-friendly way of pest management. Plant Blumea eriantha (DC), after identification and authentication from Botanical Survey of India, Allahabad was collected, shade dried and pulverized to get powder of 50-60 mesh size, which was then extracted in distilled water with 2.63% yield through soxhletion. Obtained extracts were fractionated by thin layer chromatography (TLC) with retention factor (RF) values, namely 0.26, 0.46, 0.64 in benzene: methanol (12:0.5) and same solvents were tried in column chromatography and purified fractions were obtained. Then distilled water extract is purified fraction of Blumea eriantha (DC) was applied on larvae of mosquitoes by solvent using water in increasing concentrations (10- 100%) and placed in each beaker separately for each concentration. It was observed that purified fraction of Blumea eriantha (DC) distilled water extract at 100% inhibitory concentration was found to be much more effective, that is 53.33±5.77% mortality as compared to the standard drug Temephos (90.00±11.55% mortality). LC50 and LC90 value were found to be 36.97% and 93.25%, respectively against larvae of mosquitoes. Thus distilled water extract of the plant, Blumea eriantha (DC) was found to be effective for the management of solvent at moderate concentration.

Keywords

Larvicides, *Blumea eriantha* (DC), *Aedes aegypti*, *Anopheles stephensi*

View Full Length Paper

IJEP 40 (12) : 1323-1331 (2020)

A Pilot-Scale Comparison Of Four Filter Media For Use In The Constructed Wetland For Treating Greywater

Anudeep Nema, K.D. Yadav and R.A. Christian

Abstract

The reuse of low strength wastewater, such as domestic greywater, is a critical tool in abating the effects of water scarcity and population growth. The principal aim of this study was to examine the efficiency of the granular filtration system in greywater treatment for various filter materials. In this context, the study focused on the pollutant removal efficiency of the vertical

filters with different materials as granite (M-1), marble (M-2), gravel (M- 3), limestone (M-4) observed. Granite, marble, gravel and limestone reduced the influent total solids (TS) by 21%, 22%, 23% and 23%; total suspended solids (TSS) by 46%, 50%, 49% and 52%; chemical oxygen demand (COD) by 39%, 31%, 41% and 34%; turbidity by 69%, 60%, 71% and 62%, respectively. However, other parameters, such as total dissolved solids (TDS), electrical conductivity (EC) and calcium increased from initial but not significantly. Overall, gravel and granite filters was observed to be the most effective in improving quality of greywater in terms of organic matter, solids and turbidity reduction and used as constructed wetland media. The reliability of these filters under high and fluctuating loads and the long-term stability of the filter materials need further study.

Keywords

Media filters, Greywater, Primary treatment

View Full Length Paper

IJEP 40 (12) : 1332-1337 (2020)

Assessment Of Groundwater Quality Of Palladam Region In Tirupur District, Tamil Nadu

S. Manimegalai and K. Kaleel Rahman

Abstract

Water is very important for life, without water our life cannot move. The major sources of water are rainfall, surface water involving rivers, lakes and groundwater involving wells and bore wells. Therefore, even countries with vast water resources could suffer from a scarcity of water in the near future. The changes in quality of groundwater respond to the variation in physical, chemical and biological environments through which it passes. Hence the present investigations were carried to analyse the quality of groundwater samples in Palladam region at Tirupur district of Tamil Nadu, India. The physico-chemical parameters of water samples, such as temperature (T), pH, specific conductance, total alkalinity (TA), total dissolved solids (TDS), total hardness (TH), dissolved phosphates, dissolved oxygen (DO), chloride (Cl-), fluoride (F-), iron (Fe), calcium (Ca), biological oxygen demand (BOD) and chemical oxygen demand (COD) have been carried out in Tirupur district. The results revealed that the values of all the samples (S1, S2, S3, S4 and S5) are below the permissible limits for turbidity, specific conductance, hardness, TDS, DO, phosphates, fluorides, iron, COD and BOD. Whereas pH, temperature, chlorides and calcium showed the values are above the permissible limits by WHO and BIS standards. Thus it concluded that rainwater harvesting is one of the solutions to

minimize the inorganic chemical concentration in groundwater. It is suggested that monitoring the groundwater quality periodically to prevent further contamination is important.

Keywords

Palladam region, Tirupur district, Groundwater, Physico-chemical parameters, WHO, BIS

View Full Length Paper

IJEP 40 (12) : 1338-1342 (2020)

A Study On Air Quality Index In Mylavaram, Andhra Pradesh

V. Ramakrishna, B. Dharani Mounika, M. Haritha, T. Saiteja and A. Sambasiva Rao

Abstract

The ambient air is getting polluted due to increased vehicular movement and day-to-day needs. This causes adverse impacts on human health ranging from respiratory problems, bronchitis, asthma and cardiovascular problems. Air quality index (AQI) is developed to convert the ambient concentration of air pollutants to a single index that will help in the assessment of air quality. The AQI of New Delhi is categorized as severe with AQI values >500 during October 2019, which is very alarming. This necessitates the increasing concern for determination of air quality index at any location. Mylavaram is a small town with a population of 21,763 in Krishna district of AP but located along NH 30 which results in heavy traffic flow. The AQI at three different locations in Mylavaram is studied during March-April 2019 and 24 hr samples are collected for PM10, PM2.5, SO2 and NO2. Results reveal that the AQI in Mylavaram is satisfactory with PM2.5 and PM10 as the critical pollutants. The AQI in Vijayawada, the nearest major city, is also analyzed from secondary data for the same period and observed that it can also be categorized as satisfactory with PM10 as critical pollutant indicating similar trends for both rural area (Mylavaram) and urban area (Vijayawada). It can be concluded that vehicular emissions and airborne dust are the major problems with respect to air quality in this region.

Keywords

Air quality index, Mylavaram, Vijayawada, Ambient air quality, PM10, PM2.5

View Full Length Paper

IJEP 40 (12) : 1343-1347 (2020)

Application Of Pressmud Compost From Sugarcane Trash For Improvement Of Soil Texture And Crop Production

Shailesh S. Deshmukh

Abstract

Composting is a biological process which includes degradation of organic material in aerobic condition. Degraded organic matter undergoes chemical and physical changes to form a useful end product with homogeneity. This product has greater importance as organic fertilizer to enhance soil health. The present study was carried out at the soil testing laboratory in Dnyaneshwar Cooperative Sugar Factory Limited, Bhenda. Today environmental pollution is serious problem. Generally, farmers burn the sugar trash which creates air pollution by increasing percentage of greenhouse gases in atmosphere. This leads to drastic changes in environment causing an environmental imbalance. Air pollution also causes respiratory diseases which become a more common thing nowadays. But if farmers used this eco-friendly method of composting, air pollution can be reduced to a considerable extent. Compost manufacturing method helps to improve soil texture and increase the yield of crops. Use of compost manure can reduce the excessive use of pesticides and chemical fertilizer. The method for preparation of compost from pressmud is affordable to farmers in which simple and inexpensive equipment are used that are easily available in urban and rural areas. In India, composting is very important for sustainable development in agriculture.

Keywords

Sugarcane trash, Pressmud, Superphosphate, Microorganisms, Compost

View Full Length Paper

IJEP 40 (12) : 1348-1353 (2020)

Water Quality Analysis Of Ulsoor Lake : Pre And Post-Ganesha Festival

Latha V., Athufa Tabassum, Pallavi Gahlawat and Shaziya Sultana

Abstract

Urban lakes are the most reliable source of water in the cities during water crisis. Due to uncontrolled and unplanned urbanization, urban lakes are in the verge of extinction. Currently, rapid deterioration of lakes is a burning issue. Present study highlights the changes in physico-chemical characteristics of Ulsoor lake water prior to and after Ganesha festival. Four different

sampling points were selected, depending on the source of inflow of pollutants. Significant increase was observed in pH, electrical conductivity, chloride, sulphate, bicarbonate, calcium and potassium concentrations after idol immersion. Contrarily, the concentration of magnesium and sodium adsorption ratio (SAR) decreased following the immersion, affecting the natural composition of the lake water. Results indicate drastic changes in physico-chemical characteristics of water and highlights the need for usage of eco-safe material for idol manufacture. In addition to this, proper management of lake to maintain the ecological homeostasis is recommended.

Keywords

Ulsoor lake, Physico-chemical characteristics, Idol immersion

View Full Length Paper

IJEP 40 (12) : 1354-1356 (2020)

Algae As A Bioindicator To Determine Water Quality Of Kosi River At Bhagalpur And Katihar Districts Of North Bihar

Rashmi Kumari and Arvind Kumar

Abstract

The Kosi river was chosen as the study area in order to generate baseline data on water quality of lower stretches. Phytoplankton as pollution indicator using Palmer's index was carried out at three sites from 2017-2018. Samples were collected on a seasonal basis (winter, pre-monsoon and post-monsoon) from Kosi river. The river is facing various anthropogenic activities, like extensive agricultural practices, sand mining and cattle wallowing. Phytoplankton is an assemblage of heterogenous microscopic forms of an aquatic ecosystem whose movement is more or less dependent upon water current. The phytoplankton is photoautotrophic organism having chlorophyll a and the primary producer in the base of the food chain of the aquatic ecosystem. Palmer's index varies between 17.33 at site-I to 21 at site-III during the year 2017 of the study period, which indicate site-I has probable organic pollution and site-III has high organic pollution and during the year 2018, it varies between 12 at site-I to 20.6 at site-II, which indicate site-I has less organic pollution and site-II has high organic pollution. Site-I has low and probable organic pollution, site-II has high organic pollution and site-III has high organic pollution during both the years. Continuous monitoring of the Kosi river is necessary.

Keywords

Phytoplankton, Kosi river, Pollution, Anthropogenic activities, Palmer's index

View Full Length Paper

← Previous Post

Next Post →

Quick Link

Home

E-Journals

Subscription

Contact Us

Account

Query Form

Your name

Your email

Query / Feedback

Submit

Contact Us

Indian Journal of Environmental Protection,
Kalpana Corporation, Kalpana Bhawan,
N 10/60 H-12, Shyama Nagar
P.O. Bajardiha, Varanasi – 221106

Phone Number : +91 8130034876
Landline No : 0542-236079

Email: kalpanacorp81@gmail.com

Websites : www.e-ijep.co.in

Copyright © 2022 Indian Journal of Environmental Protection | Kalpana Corporation

Research Journal of Pharmacy and Technology
(Home.aspx)

ISSN
0974-360X (Online)
0974-3618 (Print)

Anti-infective potential of marine actinobacteria against carbapenem resistant Klebsiella pneumoniae ATCC13882 (AbstractView.aspx?PID=2020-13-8-19) (https://scholar.google.co.in/scholar?q=Anti-infective potential of marine actinobacteria against carbapenem resistant Klebsiella pneumoniae ATCC13882)

Author(s): Sukanth kumar Enmozhi (search.aspx?key=Sukanth kumar Enmozhi), Vijayalakshmi Ganesan (search.aspx?key=Vijayalakshmi Ganesan), Manigundan Kaari (search.aspx?key=Manigundan Kaari), Bharathi Selvaraj (search.aspx?key=Bharathi Selvaraj), Gopikrishnan Venugopal (search.aspx?key=Gopikrishnan Venugopal), Jerrine Joseph (search.aspx?key=Jerrine Joseph), Radhakrishnan Manikkam (search.aspx?key=Radhakrishnan Manikkam)

Email(s): mrkactinos@gmail.com (mailto:mrkactinos@gmail.com)

DOI: 10.5958/0974-360X.2020.00646.0 (https://doi.org/10.5958/0974-360X.2020.00646.0)
(https://scholar.google.co.in/scholar?q=10.5958/0974-360X.2020.00646.0)

Address: Sukanth kumar Enmozhi1, Vijayalakshmi Ganesan2, Manigundan Kaari2, Bharathi Selvaraj3, Gopikrishnan Venugopal2, Jerrine Joseph2, Radhakrishnan Manikkam2
1School of Biotechnology and Genetic Engineering, Bharadhidasan University, Tiruchirappalli.
2Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai.
3Department of Microbiology, Sri Sankara Arts and Science College, Kanchipuram.
*Corresponding Author

Published In: **Volume -** 13, **Issue -** 8, **Year -** 2020 **(Issues.aspx?VID=13&IID=8)**



(https://badge.dimensions.ai/details/doi/10.5958/0974-360X.2020.00646.0?domain=https://rjptonline.org)

View PDF View HTML (HTMLPaper.aspx?Journal=Research Journal of Pharmacy and Technology;PID=2020-13-8-19)

ABSTRACT:
Isolation of actinobacteria from mangrove rhizosphere sediments has been investigated for their antimicrobial properties. Twenty eight actinobacterial strains were isolated from the rhizosphere sediments collected from Parangipettai, Tamil Nadu and Andaman mangrove ecosystems. Strain PPA12 from Parangipettai mangrove sediments, which showed promising antimicrobial activity in agar plug method, had been selected as potential strain. The PPA12 actinobacteria produced bioactive metabolite only by agar surface fermentation and not by submerged fermentation. The bioactive metabolite from the potent strain PPA12 that was produced by agar surface fermentation which was extracted using different solvents and the maximum extraction was with methanol. The minimum inhibitory concentration (MIC) of the methanol extract was found to be 125µg/ml against carbapenem resistant Klebsiella pneumoniae. Partial purification of the crude methanol extract was done by TLC and the bioautography showed that the active fractions with RF value of 0.288 and 0.55 have bioactivity. The physiochemical, molecular characterization and phylogenetic analysis of the potent strain showed that the actinobacteria PPA12 was found to be Streptomyces fumigatiscleroticus.

Keywords: Antibiotic () Carbapenem () Actinobacteria () Antimicrobial. ()

Cite this article:
Sukanth kumar Enmozhi, Vijayalakshmi Ganesan, Manigundan Kaari, Bharathi Selvaraj, Gopikrishnan Venugopal, Jerrine Joseph, Radhakrishnan Manikkam. Anti-infective potential of marine actinobacteria against carbapenem resistant Klebsiella pneumoniae ATCC13882. Research J. Pharm. and Tech. 2020; 13(8):3653-3660. doi: 10.5958/0974-360X.2020.00646.0

Cite(Electronic):
Sukanth kumar Enmozhi, Vijayalakshmi Ganesan, Manigundan Kaari, Bharathi Selvaraj, Gopikrishnan Venugopal, Jerrine Joseph, Radhakrishnan Manikkam. Anti-infective potential of marine actinobacteria against carbapenem resistant Klebsiella pneumoniae

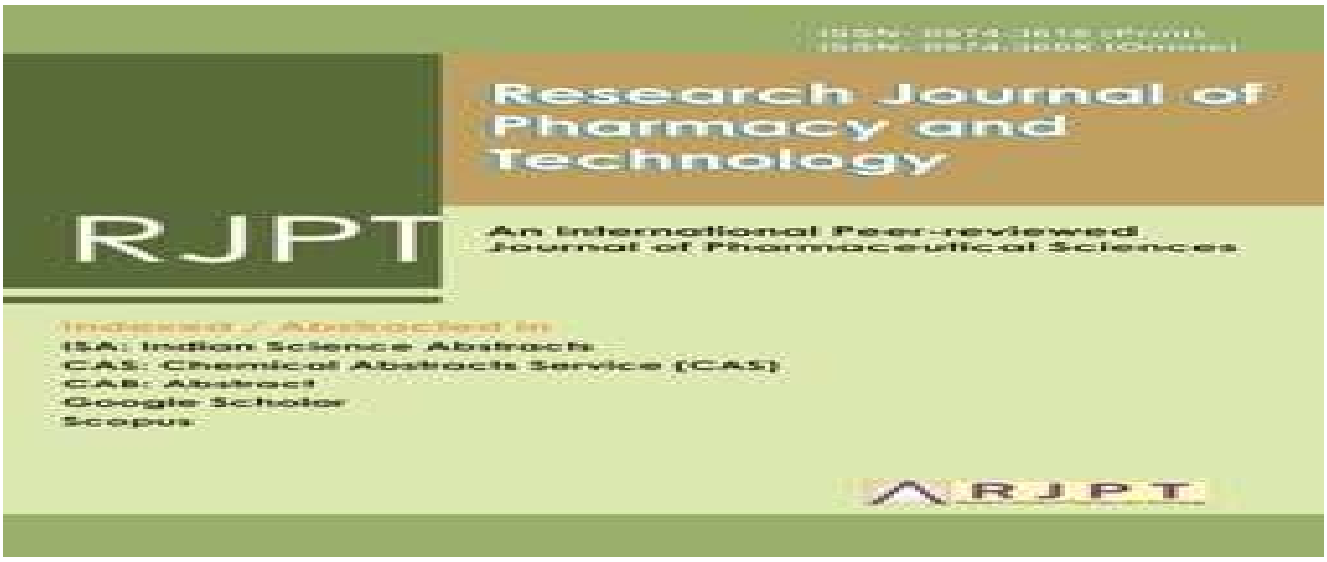
ATCC13882. Research J. Pharm. and Tech. 2020; 13(8):3653-3660. doi: 10.5958/0974-360X.2020.00646.0 Available on: https://rjptonline.org/AbstractView.aspx?PID=2020-13-8-19

REFERENCES:

1. Ibrahim AH, Samar YD, Mostafa AF, Mohamed SK, Tobias AMG and Abdelmohsen UR. Natural Product Potential of the Genus Nocardiopsis. Marine Drugs. 2018; 16(5): 147.
2. Manivasagan P, Kyong-Hwa K, Sivakumar KECY, Li-Chan, Hyun-Myung Oh and Se-Kwon Kim. Marine actinobacteria: An important source of bioactive natural products. Environ. Toxicol. Pharmacol. 2014; 38(1):172-188.
3. Yu J, Liu Z, Qiu L, Xiaohui Q, Ying J and Beom SK. Isolation and characterization of actinobacteria from Yalujiang coastal wetland, North China. Asian Pacific J Tropical Biomed. 2015; 5(7):555-560.
4. Elbendary AA, Hessain AM, El-Hariri MD, Seida AA, Moussa IM, Mubarak AS, Kabli SA, Hemeg HA and El Jakee JK. Isolation of antimicrobial producing Actinobacteria from soil samples. Saudi J. Biol. Sci. 2018; 25(1):44-46.
5. Meletis G. Carbapenem resistance: overview of the problem and future perspectives. Therapeutic adv. Infectious Dis. 2016; 3(1):15-21.
6. Sathiyaseelan K and Stella D. Isolation, identification and antimicrobial activity of marine actinomycetes isolated from parangipettai. Rec. Res. Sci. Tech. 2011; 3(9): 74-77.
7. Radhakrishnan M, Balaji S and Balagurunathan R. Thermotolerant actinomycetes from the Himalayan Mountain-Antagonistic potential, characterization and identification of selected strains. Malaysian Appl. Bio. J. 2007; 36(1): 59-65.
8. Mohamed H, Miloud B, Zohra F, Garcia-Arenzana JM, Veloso A and Rodriguez-Couto S. Isolation and Characterization of actinobacteria from Algerian Sahara soils with antimicrobial activities. Int. J. Mol. Cellular Med. 2017; 6(2):109-120.
9. Radhakrishnan M, Pazhanimurugan R, Gopikrishnan V, Balagurunathan R and Vanaja Kumar. Streptomyces sp D 25 isolated from Thar Desert soil, Rajasthan producing pigmented antituberculosis compound only in solid culture. J. Pure Appl. Microbiol. 2014; 8(1): 333-337.
10. Sathish KS and Kokati VBR. In-vitro antimicrobial activity of marine actinobacteria against multidrug resistance Staphylococcus aureus. Asian Pacific J. Tropical Biomed. 2012; 2(10):787-792.
11. Eccleston GP, Brooks PR and Kurtboke DI. The occurrence of bioactive Micromonosporae in aquatic habitats of the sunshine coast in Australia. Marine Drug. 2008; 6243-6261.
12. Bavya M, Mohanapriya P, Pazhani murugan R and Balagurunathan R. Potential bioactive compounds from marine actinomycetes against biofouling bacteria. Indian J. Geomarine Sci. 2011; 40(4): 578-582.
13. Selvameenal L, Radakrishnan M and Balagurunathan R. Antibiotic pigment from desert soil actinomycetes; biological activity, purification and chemical screening. Indian J. Pharma Sci. 2009; 71: 499 - 504.
14. Augustine SK, Bhavsar SP and Kapadnis BP. A non-polyene antifungal antibiotic from Streptomyces albidoflavus PU23. J. Biosci. 2005; 30: 201 - 211.
15. Balagurunathan R and Radhakrishnan M. Biotechnological, genetic engineering and nanotechnological potential of actinomycetes. 2010; p. 436 In D.K. Maheshwari, R.C. Dubey, and R. Saravanamuthu (ed.), Industrial Exploitation of Microorganisms. I.K. International Publishing House Pvt. Ltd, New Delhi.
16. Shirling EB and Gottileb D. Methods for characterization of Streptomyces species. Int. J. Systematic Bacteriol. 1966; 6(3): 313 -340.
17. Gothwal RK, Nigam VK, Mohan MK, Sasmal D and Ghosh P. Screening of nitrogen fixers from rhizospheric bacterial isolates associated with important desert plants. Appl. Ecol. Environ Res. 2008; 6(2): 101-109.
18. Takahashi YO and Omura S. Isolation of new actinomycete strains for the screening of new bioactive compounds. J Gen Applied Microbiol. 2003; 49: 141-154.
19. Thompson JD, Higgins DG and Gibson TJ. Clustal W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22: 4673-4680.
20. Boudjella H, Bouti K, Zitouni A, Mathieu F, Lebrihi A and Sabaou N. Taxonomy and chemical characterization of antibiotics of Streptosporangium Sg 10 isolated from a Saharan soil. Microbiol. Res. 2006; 161: 288-298.

21. Felsenstein J. Phylogenies and the Comparative Method. The American Naturalist. 1985; 125(1): 1-15.
22. Berdy J. Thoughts and facts about antibiotics: where we are now and where we are heading. J. Antibiot. (Tokyo). 2012; 65(8): 385-395.
23. Tiwari, K. and Gupta, R.K. 2012. Rare actinomycetes: a potential storehouse for novel antibiotics. Crit. Rev. Biotechnol. 32(2):108-32.
24. Radhakrishnan M, Vijayalakshmi G, Gopikrishnan V and Jerrine J. Bioactive potential of actinobacteria isolated from certain understudied regions in India. J. Applied Pharma. Sci. 2016; 6(8): 151-155.
25. Mohana S and Radhakrishnan M. Streptomyces sp MA7 isolated from mangrove rhizosphere sediment effective against Gram negative bacterial pathogens. International Journal of PharmTech Research. 2014; 6(4): 1259-1264.
26. Diraviyam T, Radhakrishnan M and Balagurunathan R. Antioxidant activity of melanin pigment from Streptomyces species D5 isolated from Desert soil, Rajasthan, India Drug Invention Today. 2010; 3(3): 12-13.
27. Knight V, Sanglier JJ, DiTullio D, Braccili S, Bonner P, Waters J, Hughes D and Zhang L. Diversifying microbial natural products for drug discovery. Appl. Microbiol. Biotechnol. 2003; 62: 446-458.
28. Salamoni SP, Mann MB, Campos FS, Franco AC, Germani JC and Van der Sand ST. Preliminary characterization of some Streptomyces species isolated from a composting process and their antimicrobial potential. World J. Microbiol. Biotechnol. 2010; 26(10): 1847-1856.
29. Lalitha C, Thiagarajan R, Sudarshan S, Rathore and Jayapradha R. 2010. Bioactive molecule from Streptomyces sp. mitigates MDR Klebsiella pneumoniae in Zebrafish infection model. Front. Microbiol. 2010; 8: 614.
30. Treskatis SK, Orgeldinger V, Wolf H and Gilles ED. Morphological characterization of filamentous microorganisms in submerged cultures by on-line digital image analysis and pattern recognition. Biotechnology and Bioengineering. 1997; 53(2):191-201.
31. Manteca A, Alvarez R, Salazar N, Yague P and Sanchez J. Mycelium differentiation and antibiotic production in submerged cultures of Streptomyces coelicolor. Applied and Environmental Microbiology. 2008; 74(12): 3877-3886.

RECOMONDED ARTICLES:



Research Journal of Pharmacy and Technology (RJPT) is an international, peer-reviewed, multidisciplinary journal....

Read more >>> (AboutJournal.aspx)

RNI: CHHENG00387/33/1/2008-TC
DOI: 10.5958/0974-360X

0.38
56th percentile
2018
CiteScore
Powered by Scopus
(https://www.scopus.com/sourceid/21100197160?dgcid=sc_widget_citescore)



(https://www.scimagojr.com/journalsearch.php?q=21100197160&tip=sid&exact=no)

QUICK LINKS

SUBMIT ARTICLE (SUBMITARTICLE.ASPX)

AUTHOR'S GUIDELINES (DOWNLOADS/INSTRUCTIONS_TO_AUTHOR.PDF)

PAPER TEMPLATE (DOWNLOADS/PAPER_TEMPLET.DOC)

COPYRIGHT FORM (DOWNLOADS/COPYRIGHT TRANSFER FORM.DOCX)

CERT. OF CONFLICT OF INTREST (DOWNLOADS/CERTIFICATE OF CONFLICT OF INTREST.PDF)

PROCESSING CHARGES (CHARGESDETAILS.ASPX)

INDEXING INFORMATION (INDEXED_IN.ASPX)

LATEST ISSUES

JANUARY 2023 (80) (ISSUES.ASPX?VID=16&IID=1)

DECEMBER 2022 (91) (ISSUES.ASPX?VID=15&IID=12)

NOVEMBER 2022 (91) (ISSUES.ASPX?VID=15&IID=11)

OCTOBER 2022 (88) (ISSUES.ASPX?VID=15&IID=10)

SEPTEMBER 2022 (87) (ISSUES.ASPX?VID=15&IID=9)

AUGUST 2022 (82) (ISSUES.ASPX?VID=15&IID=8)

JULY 2022 (80) (ISSUES.ASPX?VID=15&IID=7)

JUNE 2022 (82) (ISSUES.ASPX?VID=15&IID=6)

POPULAR ARTICLES

(AbstractView.aspx?PID=2017-10-9-42)
Detection of Food Adulterants in Chilli, Turmeric and Coriander Powders by Physical and Chemical Methods
(AbstractView.aspx?PID=2017-10-9-42)

(AbstractView.aspx?PID=2020-13-1-43)
Formulation and Evaluation of Herbal Face Cream
(AbstractView.aspx?PID=2020-13-1-43)

(AbstractView.aspx?PID=2013-6-2-15)
Medicinal Plants from Solanaceae Family
(AbstractView.aspx?PID=2013-6-2-15)

(AbstractView.aspx?PID=2020-13-4-16)
Formulation and Evaluation of Herbal Lipsticks

(AbstractView.aspx?PID=2020-13-4-16)

(AbstractView.aspx?PID=2020-13-7-74)
Pharmaceutical Incompatibilities: Causes, Types and Major ways of Overcoming in Extemporaneous Medicinal forms

(AbstractView.aspx?PID=2020-13-7-74)

(AbstractView.aspx?PID=2019-12-1-69)
Recent Advances in Preventive Resin Restoration (PRR)

(AbstractView.aspx?PID=2019-12-1-69)

(AbstractView.aspx?PID=2017-10-9-19)
Formulation and Evaluation of Aspirin Tablets by Using Different Lubricants in Combination for better Kinetic Drug Release Study by PCP

(AbstractView.aspx?PID=2017-10-9-19)

(AbstractView.aspx?PID=2014-7-9-14)
The Use of Neem in Oral Health

(AbstractView.aspx?PID=2014-7-9-14)

(AbstractView.aspx?PID=2018-11-2-70)
Recent Advancements in Laminates and Veneers in Dentistry

(AbstractView.aspx?PID=2018-11-2-70)

(AbstractView.aspx?PID=2019-12-11-80)
Dental Waxes–A Review

(AbstractView.aspx?PID=2019-12-11-80)

(AbstractView.aspx?PID=2017-10-12-61)
Mathematical Models in Drug Discovery, Development and Treatment of Various Diseases – A Case Study

(AbstractView.aspx?PID=2017-10-12-61)

(AbstractView.aspx?PID=2010-3-3-60)
Evaluation of Ayurvedic Marketed Formulations Asava's and Arista's.

(AbstractView.aspx?PID=2010-3-3-60)

(AbstractView.aspx?PID=2014-7-9-16)
Project Writing for Retail Pharmacy Practical Training: A Proforma

(AbstractView.aspx?PID=2014-7-9-16)

(AbstractView.aspx?PID=2012-5-5-5)
A comparative, Bioequivalence study to evaluate the safety and pharmacokinetic profile of single dose Ivabradine 7.5mg Tablets in healthy, adult, human subjects under fasting condition.

(AbstractView.aspx?PID=2012-5-5-5)

(AbstractView.aspx?PID=2020-13-3-81)
Regulatory requirements for conducting Clinical Trials in India

(AbstractView.aspx?PID=2020-13-3-81)

Recent Articles

Tags

Not Available

ABOUT JOURNAL

Research Journal of Pharmacy and Technology (RJPT) is an international, peer-reviewed, multidisciplinary journal, devoted to pharmaceutical sciences. The aim of RJPT is to increase the impact of pharmaceutical research both in academia and industry, with strong emphasis on quality and originality. RJPT publishes Original Research Articles, Short Communications, Review Articles in all areas of pharmaceutical sciences from the discovery of a drug up to clinical evaluation. Topics covered are: Pharmaceutics and Pharmacokinetics; Pharmaceutical chemistry including medicinal and analytical chemistry; Pharmacognosy including herbal products standardization and Phytochemistry; Pharmacology: Allied sciences including drug regulatory affairs, Pharmaceutical Marketing, Pharmaceutical Microbiology, Pharmaceutical biochemistry, Pharmaceutical Education and Hospital Pharmacy.
Read More >>> (AboutJournal.aspx)

VISITORS



Today: 23324

Yesterday: 24471

Total: 20343751

(https://tlabssolutions.com/) Designed and Developed by:
T-Labs Solutions (https://tlabssolutions.com/)

ISSN 0974-3618 (Print)
0974-360X (Online)

www.rjptonline.org

RESEARCH ARTICLE

Anti-infective potential of marine actinobacteria against carbapenem resistant *Klebsiella pneumoniae* ATCC13882

Sukanth kumar Enmozhi[1], Vijayalakshmi Ganesan[2], Manigundan Kaari[2], Bharathi Selvaraj[3], Gopikrishnan Venugopal[2], Jerrine Joseph[2], Radhakrishnan Manikkam[2]

[1]School of Biotechnology and Genetic Engineering, Bharadhidasan University, Tiruchirappalli
[2]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai
[3]Department of Microbiology, Sri Sankara Arts and Science College, Kanchipuram
*Corresponding Author E-mail: **mrkactinos@gmail.com**

ABSTRACT:

Isolation of actinobacteria from mangrove rhizosphere sediments has been investigated for their antimicrobial properties. Twenty eight actinobacterial strains were isolated from the rhizosphere sediments collected from Parangipettai, Tamil Nadu and Andaman mangrove ecosystems. Strain PPA12 from Parangipettai mangrove sediments, which showed promising antimicrobial activity in agar plug method, had been selected as potential strain. The PPA12 actinobacteria produced bioactive metabolite only by agar surface fermentation and not by submerged fermentation. The bioactive metabolite from the potent strain PPA12 that was produced by agar surface fermentation which was extracted using different solvents and the maximum extraction was with methanol. The minimum inhibitory concentration (MIC) of the methanol extract was found to be 125µg/ml against carbapenem resistant *Klebsiella pneumoniae*. Partial purification of the crude methanol extract was done by TLC and the bioautography showed that the active fractions with RF value of 0.288 and 0.55 have bioactivity. The physiochemical, molecular characterization and phylogenetic analysis of the potent strain showed that the actinobacteria PPA12 was found to be *Streptomyces fumigatiscleroticus.*

KEYWORDS: Antibiotic, Carbapenem, Actinobacteria, Antimicrobial.

INTRODUCTION:

Marine environment is largely an untapped source for obtaining actinobacteria, having potential to produce novel, bioactive natural products. Actinobacteria are the prolific producers of pharmaceutically active secondary metabolites, accounting for about 70% of the naturally derived compounds that are currently in clinical use[1]. Among the various actinobacterial genera, *Actinomadura, Actinoplanes, Amycolatopsis, Marinispora, Micromonospora, Nocardiopsis, Saccharopolyspora, Salinispora, Streptomyces* and *Verrucosispora* are the major potential producers of commercially important bioactive natural products.

In this respect, *Streptomyces* ranks first with a large number of bioactive natural products. Marine actinobacteria are unique with diverse biological properties including antimicrobial, anticancer, antiviral, insecticidal and enzyme inhibitory activities. They have attracted global attention in the last ten years for their ability to produce pharmaceutically active compounds[2]. The isolation rates of antibiotic producing actinobacteria from marine soil sediments were approximately 27%. Many studies on the antimicrobial activity of the actinobacteria isolated from marine environment have been done, but the results are different. It was reported that the components of the medium, samples differ in types and the exact reason could influence the antimicrobial activities of actinobacteria[3].

Since the discovery of antibiotics, bacterial resistance to antibiotic has continued to evolve. Thus, we are witnessing more and more multi-resistant bacteria that pose a serious public health problem[4] and carbapenem resistance also has been major threat. It is reasonable

Received on 10.09.2019 Modified on 15.11.2019
Accepted on 20.01.2020 © RJPT All right reserved
Research J. Pharm. and Tech. 2020; 13(8):3653-3660.
DOI: 10.5958/0974-360X.2020.00646.0

however to consider that starting from the emergence of carbapenem resistance until its widespread dissemination, some human-related factors play a crucial role[5]. So, the hunt for new production of antibiotics is essential. Due to these reasons the present study aims on the production of antibiotics through the marine soil sediment of Parangipettai and Andaman coast line which has been evinced to have high activity by producing enzymes, antitumour agents and antibiotics[6].

MATERIALS AND METHODS:

Sample collection and pretreatment

The soil samples were collected from Parangipettai Marshlands (11.5084° N, 79.7568° E) and Andaman Island coastal region (11.7401° N, 92.6586° E). About 15 g of the soil sample was weighed and dried at room temperature for 3 days. The air dried samples was kept in hot air oven at 55°C for 15 minutes[7].

Isolation of actinobacteria:

Actinobacteria were isolated by adopting standard spread plate method using two different media such as Starch casein agar (SCA) Starch - 10g, Caesin - 0.3g, K_2NO_3 - 2.0g, $MgSO_4$ - 0.05g, $FeSO_4$ - 0.01g, $CaCO_3$ - 0.02g, NaCl - 2.0g, distilled water -1000ml, Agar – 22g and Kuster's agar (KUA) Glycerol-10 g, Caesin - 0.3g, KNO_3 - 0.2g, K_2HPO_4 - 2g, Soluble Starch- 0.5g, Asparagine-0.1g, $FeSO_4.7H_2O$ - 0.01g, $CaCO_3$ - 0.02g, $MgSO_4$ - 0.05g, distilled water - 1000ml, Agar – 22g. All the media used in this study were prepared using 50% sea water[4]. Colonies showing suspected actinobacterial morphology were recovered from the isolation plates using sterile L-shaped needle and inoculated on yeast extract malt extract agar (YEME) Yeast Extract-4g, Malt Extract - 10g, Glucose - 4g, NaCl - 10g, Agar - 22g. After incubation at 28^0C for 7 days, morphologically different actinobacterial cultures were selected and preserved using YEME agar slants as well as in 30% glycerol[8].

Production of bioactive compounds:

Bioactive compound from actinobacterial cultures was produced by agar surface fermentation[9]. All the actinobacterial cultures were inoculated into yeast extract malt extract (YEME) agar plates and incubated at 28^0C for 10 days for the production of secondary metabolites. During incubation, the extracellular metabolites are secreted into the agar medium.

Testing for antimicrobial activity:

Antimicrobial activity of actinobacterial cultures was tested by adopting agar plug method[9]. Actinobacterial cultures grown on YEME agar plates were taken and the mycelial growth was removed from the agar surface using sterile spatula. Test human pathogens such as *Staphylococcus aureus* (ATCC 29213), *Escherichia coli* (ATCC 13439), *Klebsiella pneumonia* (ATCC 13882), *Pseudomonas species (*ATCC 27853*)*, *Candida albicans* (ATCC 90028) and fish pathogens *Aeromonas hydrophilla*, *Vibrio species*, *Vibrio harveii* were used. All bacterial pathogens with 0.5 McFarland standards were inoculated into nutrient agar plates and *Candida albicans* was inoculated on Sabouraud dextrose agar, using sterile cotton swab. Agar plug with 5 mm diameter were cut from the YEME agar grown with actinobacterial cultures were placed over agar seeded with test pathogens. All the plates were incubated at 37^0C for 24 hours and observed for zone of inhibition[10]. The Actinobacterial strain that showed maximum zone of inhibition against test pathogens was selected as the potential strain for further studies.

Effect of fermentation method on bioactive compound production:

Effect of solid state and submerged fermentation on bioactive compound production by the strain PPA12 was investigated. Spores of the actinobacterial strain PPA12 was inoculated into five YEME agar plates and 100 ml of YEME broth. YEME agar plates were incubated at 28^0C for 12 days. YEME broth containing flasks were incubated in rotary shaker with 95 rpm for 12 days. For every 24 hours, agar plug from YEME agar plates were taken and tested against Carbapenem resistant *Klebsiella pneumonia*[11]. Each 2ml of YEME broth from was taken and centrifuged at 10, 000rpm for 10 minutes. The cell free supernatant also tested against CRKP by adopting well diffusion method[12].

Extraction of bioactive compounds:

The actinobacterial strain produced the bioactive compounds only in solid culture and not in submerged fermentation. Hence, further production of bioactive compound from strain PPA12 was carried out by agar surface fermentation. After incubation, mycelial growth was removed aseptically using sterile spatula. The agar medium was cut into pieces and extracted overnight at room temperature using different solvents (1:2 ratio) such as methanol, chloroform, ethyl acetate and n-hexane. The solvent extracts were concentrated using rotary evaporator and quantified. Antimicrobial activity of different solvent extracts was tested against *K. pneumoniae* by disc diffusion method[13]. Zone of inhibition was measured after 24 hours of incubation at 37°C and expressed in millimetre in diameter. Solvent extract which showed maximum activity against the test pathogen was used to extract large quantity of bioactive compounds from 1000 ml of YEME agar medium.

Minimum Inhibitory Concentration of crude extract against K. Pneumoniae*:*

The MIC of the crude extract against *K. pneumoniae* was determined using microbroth dilution assay as described

in the CLSI protocols. The overnight *K. pneumoniae* culture was diluted to obtain the final concentration of 10^5 CFU/ml. The compound was dissolved in PBS and was tested at concentrations from 8 to 250 in twofold step intervals. Each well contains 800 µl of nutrient broth, 100µl of culture and 100µl of compound with varying concentrations. Culture control and medium control were maintained. All the tubes were incubated at 37°C for 18-24 hours and then observed for growth inhibition. The lowest antibiotic concentration which showed low turbidity was taken as minimal inhibitory concentration (MIC).

Partial purification of Bioactive Compound:

Analytical Thin Layer Chromatography:

Methanol extract of PPA12 crude extract was spotted on the bottom of TLC sheet with glass capillary tube and marked as solvent front. Chromatography was done with different solvents like chloroform, methanol, ethyl acetate, acetone, dichloromethane and butanol in different ratio. The separated spots were visualized with naked eyes, iodine vapours, UV light, vanillin stain and potassium permanganate strain. Rf value of the separated spots were calculated by standard formula[14].

Bioautogram:

Developed chromatogram sheets were subjected to bioautogram. 1% tetrazolium salt solution was prepared and stored in amber colored bottle. 2ml of the stock tetrazolium salt solution and 100µl of 0.5 Mac Farland's standard *K. pneumoniae* were added to molten nutrient agar at hand bearable temperature with constant stirring and poured into petriplates. The developed chromatogram was placed on the agar surface in such a way that the silica faces the medium. The plates were incubated overnight and observed for results[13].

Characterization and generic identification of potential Actinobacterial strain.

Micromorphology:

Micromorphological characteristics of potential strains were studied by adopting slide culture method. The recorded microscopic characteristics include the presence/absence of aerial mycelium, substrate mycelium, mycelial fragmentation and the details of spore chain morphology. Spore structure and spore surface details were recorded using light microscope[15].

Cultural characteristics were studied by inoculating the growth of actinobacteria strain PPA12 into different ISP (International Streptomyces Project) media such as tryptone agar (ISP1), yeast extract-malt extract agar (ISP2), oatmeal agar (ISP3), inorganic salts-starch agar (ISP4), glycerol-asparagine agar (ISP5), peptone-yeast extract-iron agar (ISP6) and tyrosine agar (ISP7). All the media were prepared by following the guidelines described by Shirling and Gottileb (1966)[16]. Plating was done in triplicate and all the plates were incubated for 10-14 days at 28°C. Cultural characteristics recorded include the nature of growth, consistency, aerial mass colour, presence of reverse side pigment and the details of soluble pigment production, if any.

Physiological Characteristics:

Basal agar medium recommended by Shirling and Gottileb (1966)[16] was used for carbon utilization studies. After sterilization of basal agar medium by autoclaving at 15 lb for 15 minutes, about 1% of various ether sterilized sugars were added and poured into petri plates. About one ml of well grown culture of actinobacterial strain PPA12 was inoculated as a straight line in different sugar containing basal agar plates. Growth was recorded after 14 days of incubation at 28°C. The same procedure was used to study the utilization of nitrogen sources by the potential actinobacteria strains. Sugars used in this study include glucose, fructose, sucrose, mannitol, inositol, xylose, rhamnose, raffinose, and cellulose (Hi media). Nitrogen sources used include aspargine, glutamine and tyrosine. Effect of different pH (5, 7, 9 and 11), temperature (20°C, 30°C, 40°C and 50°C), sodium chloride (0%, 1, 2.5, 5.0, 7.5 and 10%) and sea water (0%-100%) on growth was determined by inoculating the actinobacterial strain PPA12 onto ISP2 agar medium with respective physiological condition[7]. All the plates were observed for growth after 10 days of incubation.

Molecular Characterization:

The molecular characterization of strain PPA12 was carried out at Eurofins Genomics India pvt. Ltd., Bangalore.

Extraction of chromosomal DNA:

The chromosomal DNA of actinobacterial strain PPA12 was extracted using solute ready® genomic DNA kit. The extracted DNA was suspended in TE buffer.

PCR amplification of 16S rRNA:

The PCR amplification of 16S ribosomal RNA gene of strain PPA12 was performed by using two primers 27F 5′-AGAGTTTGATCMTGGCTCAG-3′ (forward) and 1492R 5′-TACGGYTACCTTGTTACGACTT-3′ (reverse)[17] using thermal cycler (Eppendorf Master cycler-nexus). The final volume of reaction mixture of 20 µl contained 50ng of template DNA (1.0µl), 10p mol (each 2µl) of primer (forward and reverse separately), 10µl PCR master mix (2µl Taq DNA polymerase, 10X Taq reaction buffer, 2mM $MgCl_2$, 1µl of 10Mm dNTPs and PCR additives) and 5.0µl of double-distilled water. The amplified *16S rRNA* gene was purified using 2% agarose gel prepared in TE buffer.

Sequencing of 16S rRNA:
Sanger dideoxy chain terminator sequencing was done in an automated DNA sequencer. In this fluorescence based method of DNA sequencing, four reaction mixers were prepared. Each containing only one ddNTP, was labeled with four different fluorescent dyes so that only one reaction mix was required. In addition, polymerase enzymes such as dATP, dGTP, dCTP, dTTP were also added. The sequence detection was made by subjecting labeled ddNTP to a single capillary tube. DNA fragments of different colours were separated by their respective sizes in a single electrophoretic gel. The sequences were read by determining the sequence of the colours emitted and the information was fed directly to a computer which determined the sequence.

16S rRNA sequence analysis:
The closely related 16S rRNA gene sequence of strain PPA12 homologs sequence were determined and compared for similarity level with the reference species of Streptomyces contained in genomic database banks, using the 'NCBI Blast' (www.ncbi.nlm.nih.gov) web site. After BLAST analysis, a dataset of potential orthologs was prepared by considering those database sequences which had > 99% sequence identity with the query sequence of strain PPA12.

Phylogenetic analysis:
Phylogenetic and molecular evolutionary analyses of potential strains PPA12 was conducted using software included in MEGA version 4[18] package. The 16S rRNA sequence of the strains PPA12 was aligned (http://www.ebi.ac.uk/clustalw)[19] against corresponding nucleotide sequences of representatives of the genus *Streptomyces* retrieved from GenBank[20]. Evolutionary distance matrices were generated as described and a phylogenetic tree was inferred by the Neighbor joining method. Tree topologies were evaluated by bootstrap analysis[21] based on 1000 resembling of the neighbor joining data set.

Nucleotide accession number:
The partial 16S rRNA nucleotide sequence of the potential actinobacteria strain PPA12 were deposited at GenBank database by using Bankit tool (http://www.ncbi.nlm.nih.gov/WebSub/?tool=genbank).

RESULTS:

Isolation and selection of actinobacterial cultures
Colonies with actinobacterial morphology was observed on both the isolation media SCA and KUA in which more number of actinobacterial colonies were observed on KUA than the SCA media. Thirty seven actinobacterial colonies were recovered from the two isolation media and subcultured on ISP2 agar plates. After incubation, about 28 morphologically different actinobacterial cultures were selected for the further study. They showed wide pattern of morphology on YEME agar as well as under microscopic observation.

Screening of selected actinobacterial cultures for antimicrobial activity
In the preliminary antimicrobial screening, 24 out of 28 actinobacterial strains showed activity against atleast one of the four bacterial pathogens tested (Table 1). In particular 13 cultures were found to be active against Carbapenam resistant strains. Based on the results of preliminary antimicrobial screening, actinobacterial strain PPA12, which showed maximum zone of inhibition against six bacteria out of eight bacterial pathogens tested, was selected as potential strain for further studies

Table 1: Antimicrobial activity of isolated actinobacteria by agar plug method

S. No	Strain No.	*S. aureus*	*E. coli*	*K. pneumoniae*	*A. hydrophilia*	*Vibrio sp*	*Pseudomonas sp*	*Vibrio harvei*	*C. albicans*
1	PPA2	-	-	19	18.5	18	-	15	-
2	PPA25	-	-	19	18	21	24	-	13
3	PPA22	-	-	25	-	-	-	25	25
4	PPA23	-	-	21	13	1	-	-	-
5	PPA20	11	-	18	14	19	25	24	26
6	PPA21	11	-	-	27	24	21	23	-
7	PPA11	11	-	18	23	23	21	22	-
8	PPA28	-	-	19	-	-	-	-	-
9	AIA6	12	-	-	-	-	-	-	-
10	PPA12	11	-	19	25	22	22	23	-
11	PPA19	11	-	-	23	22	17	25	-
12	PPA24	-	-	-	-	-	-	-	-
13	PPA17	19	-	-	21	22	21	22	-
14	PPA16	-	-	19	18	17	-	21	-
15	PPA13	19	-	-	19	22	18	25	-
16	PPA26	13	10	-	-	7	-	22	-
17	PPA7	-	-	12	18	21	34	22	-
18	PPA10	10	-	-	15	19	19	20	-

19	PPA14	-	-	19	16	23	-	24	-
20	PPA13	-	-	19	-	19	-	19	-
21	PPA45		-	-	-	-	-	-	-
22	PPA43		11	19	20	-	10	19	-
23	PPA15		11	-	-	13	-	-	-
24	PPA35		-	-	-	-	-	-	-
25	PPA26		-	-	-	-	-	-	-
26	PPA15		10	-	-	13	-	-	-
27	PPA28		-	-	-	-	-	-	-
28	PPA42		-	12	-	-	-	12	-

Effect of fermentation method on bioactive metabolite production:

Effect of fermentation method on bioactive compound production by the strain PPA12 is given in Table 2. Even after 10 days, the cell free supernatant was not showed activity whereas the agar plug from YEME agar showed activity from 4th day of fermentation. Hence further bioactive compound production was carried out by agar surface fermentation.

Table 2: Effect of fermentation media on bioactive metabolite production

Incubation period	Submerged fermentation In YEME broth		Solid state fermentation YEME agar	
	Growth	Zone of inhibition (mm) against *K. pneumoniae*	Growth	Zone of inhibition (mm) against *K. pneumoniae*
DAY1	+	-	+	-
DAY2	+	-	+	-
DAY3	++	-	++	-
DAY4	++	-	++	20
DAY5	++	-	++	21
DAY6	++	-	++	21
DAY7	++	-	++	22
DAY8	++	-	++	23
DAY9	+++	-	++	23
DAY10	+++	-	++	24

+++ - good growth; ++ - moderate growth; + - poor growth

Extraction of bioactive compound:

Among the four different solvents tested for extraction, methanol extract showed maximum antimicrobial activity against *K. pneumoniae*. Hence further large scale extraction was done using methanol.

Determination of MIC:

Minimal Inhibitory Concentration (MIC) of the aqueous methanol fraction against *K. pneumonia* was found to be 125µg/ml.

Partial purification of bioactive metabolite:

In TLC analysis, among the different solvent systems tested, the methanol extract of strain PPA12 showed three well separated spots viz., PPA12 - A (0.288), PPA12 - B (0.403) and PPA12 - C (0.55) when the Hexane : Ethyl acetate: Methanol used as solvent system at 6:3:1 ratio.

Bioautogram:

The developed chromatogram was run bioautogram with 1% tetrasolium salt. The results showed that the zone of inhibition with PPA12-A spot as well as PPA12-C.

Characterization and taxonomy of potential Actinobacterial strain:

Phenotypic characteristics:

Under bright field microscopic observation, the vegetative substrate mycelium was lengthy and the reproductive aerial mycelium was dark and appeared in spiral (S) arrangement. The aerial and substrate mycelia did not exhibit fragmentation. Smooth, oval shaped spores were borne in long straight chains to spirals with 30-40 spores in a filament.

Strain PPA12 showed good growth on ISP2, ISP4 and ISP7 medium while moderate growth was observed on ISP1, ISP3, ISP5 and ISP6 medium. It produced powdery colonies on all the ISP media. Growth pattern of actinobacterial strain PPA12 is as shown in table 3.

Physiological characteristics:

Strain PPA12 was able to grow well utilizing wide range of sugars except on basal medium supplemented with sucrose, raffinose, cellulose and tyrosine. Growth of strain PPA12 was observed at a wide range of temperature (20^0C – 40^0C), pH 7 and 9 and in the presence of 0 – 7.5% NaCl and 100% seawater (Table 4).

Molecular characterization:

The PCR amplification of 16s rRNA gene of actinobacterial strain PPA12 yielded around 1085 base pair sequence. The BLAST analysis showed 99% similarity to *Streptomyces fumigatiscleroticus NRRL B-3856*. Phylogenetic relationship of the strain PPA12 and related taxa are given in Figure 1. Based on the criteria given by Bosshard *et al.,* (2003), the strain PPA12 was identified to be species of the genus *Streptomyces*. The 16s rRNA gene sequence of strain PPA12 is submitted to GenBank with the accession number MK573175.

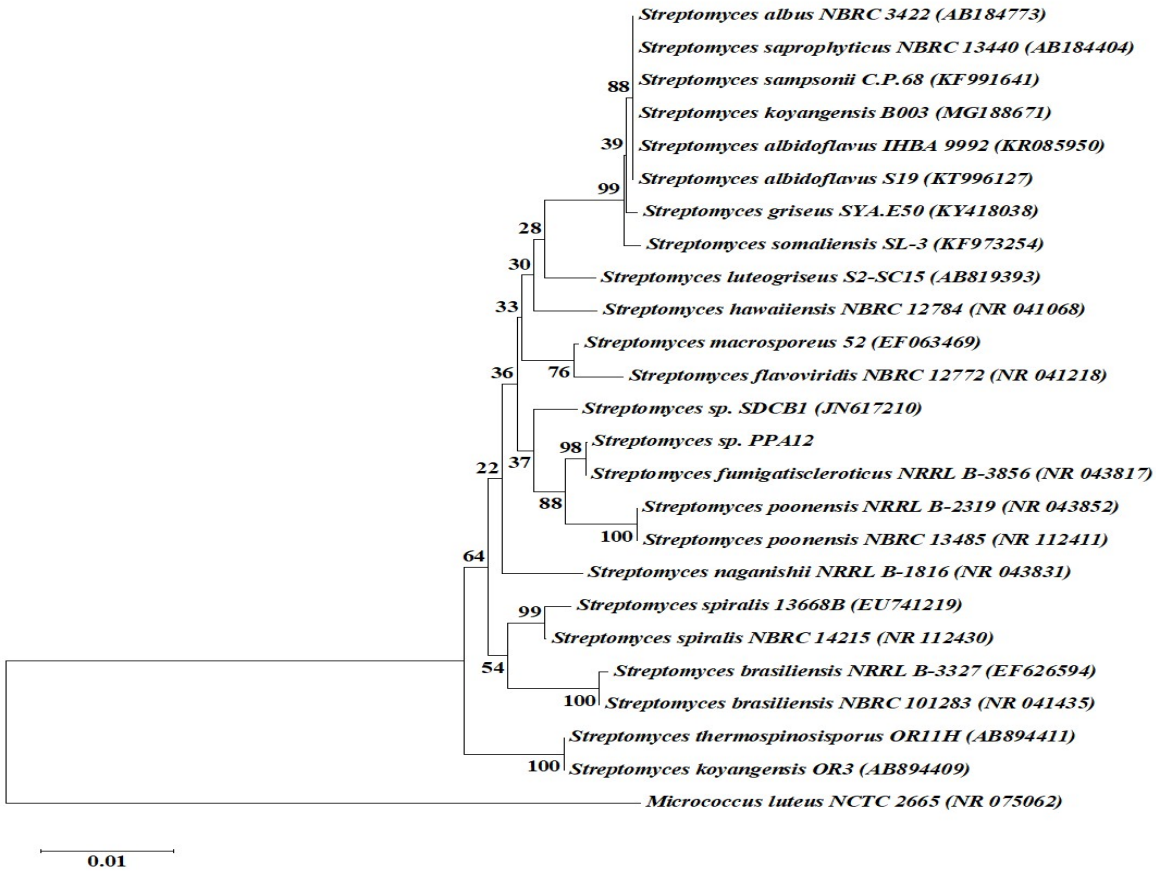


Figure 1: Phylogenetic relationship of the actinobacterial strain PPA12 and related taxa, based on 16s rRNA analysis.

Table 3: Growth pattern of actinobacteria PPA12 on different ISP agar

Medium	Growth	Colony consistency	Aerial mass colour	Reverse side pigment	Soluble pigment
Tryptone yeast extract agar (ISP1 medium)	Moderate	Powdery	Gray	-	-
Yeast extract malt extract agar (ISP 2 medium)	Good	Powdery	Gray	Golden Yellow	-
Oat Meal agar (ISP 3 medium)	Moderate	Powdery	Gray	Golden Yellow	-
Inorganic salts starch agar (ISP4 medium)	Good	Powdery	Ash	-	-
Glycerol asparagine agar (ISP 5 medium)	Moderate	Powdery	White	-	-
Peptone yeast extract iron agar (ISP6 medium)	Moderate	Powdery	White	-	-
Tyrosine agar (ISP 7 medium)	Good	Powdery	Gray	-	-

Table 4: Physiological characteristics of actinobacterial strain PPA12

Factors	Variables	Growth
Sugars	Glucose	Moderate
	Arabinose	Good
	Sucrose	Poor
	Xylose	Good
	Inositol	Good
	Mannitol	Good
	Fructose	Good
	Rhamnose	Good
	Raffinose	Poor
	Cellulose	No Growth
Amino acid	Asparagine	Good
	Glutamine	Good
	Tyrosine	Poor
Ph	5	No Growth
	7	Good
	9	Good
	11	Moderate
Temperature (^{0}C)	20	Moderate
	30	Good
	40	Moderate
	50	No Growth
NaCl %	0	Good
	1	Good
	2.5	Good
	5	Moderate
	7.5	Poor
	10	No Growth
Seawater (%)	0	Moderate
	10	Good
	25	Good
	50	Good
	100	Moderate

DISCUSSION:

As the burden of MDR *K. pneumoniae* is growing rapidly, discovery of new drugs is an indispensable process to combat those emerging resistant pathogens[22]. Actinobacteria are still the promising candidates for the discovery of novel antibiotics which are being isolated from wide range of natural habitats[23]. However, research on actinobacteria from parangipetttai mangrove ecosystems is less investigated[24]. The present study was under taken for bioprospecting of actinobacteria from parangipetttai and Andaman Island for bioactive metabolite production. In the present study, Parangipettai mangrove and Andaman island coastal region sediment were used for isolation of Actinobacteria after dry heat treatment at 55°C for 10 minutes. The use of two different media results in the isolation of more powdery or leathery colonies with substrate and aerial mycelium with the maximum no of actinobacterial colonies in Kusters agar than SCA. Mohana and Radhakrishnan (2014)[25] also recommended Kusters agar as a suitable medium for the recovery of actinobacteria from mangrove ecosystem. In the present study majority of the recovered actinobacterial cultures are *Streptomyces*. Most of the isolated actinobacterial cultures showed powdery growth with aerial and substrate mycelium. These cultures may come under the genus *Streptomyces*[16]. *Streptomyces species* are reported to produce wide range of pigments with different biological activities[26]. Further the heat treatment method and antibacterial and antifungal antibiotic supplemented in the isolation medium resulted in the reduction of unwanted bacterial and fungal growth, respectively.

According to Berdy (2012)[22] great numbers of antibiotic compounds exhibit exclusive activities against gram positive bacteria while only 1.5% are active against gram negative bacteria. Eco-physiological conditions of particular ecosystem greatly influence the biological and metabolic activity and diversity of actinobacteria[27]. Salamoni *et al.,* (2012)[28] reported that 25 *Streptomyces* species isolated from compost showed 20 different patterns of activity against 53 bacterial and fungal pathogens. Lalitha *et al.,* (2017)[29] isolated 51 morphologically different Actinobacteria isolated from rhizosphere soil of five different medicinal plants. However, only one isolate was shown to have a potential activity against MDR *K. pneumoniae*. But in the present study nearly 50% of the isolated actinobacteria showed activity against CRKP.

Strain PPA12 showed good growth on ISP2 broth during submerged fermentation by shake flask method. But, in well diffusion test, the cell free supernatant showed no activity against the test organisms. However, in agar surface fermentation, strain PPA12 showed good growth and the methanol extract obtained from the fermented medium showed activity against carbapenam resistant *K. pneumoniae.* Production of a majority of industrially important secondary metabolites from actinobacteria is carried out using submerged fermentation processes where they exhibit diverse morphological forms[30]. Morphology is influenced by environmental conditions such as medium composition and shear stress[31]. Further optimization of medium component and concentration may be needed to produce the compound from strain PPA12 by submerged fermentation.

Majority of the actinobacterial metabolites are extracellular in nature and they are extracted using the medium polar solvent ethyl acetate[7,13]. In the present the only the extracellular metabolites extracted in methanol showed antimicrobial activity. This result indicates the highly polar nature of the active compound produced by the actinobacterial strain PPA12. The methanol fraction of strain PPA12 showed promising activity against *K. pneumoniae*. This observation paves the way for the exploration of the actinobacteiral strain PPA12 for clinical applications. The analysis of 16s rRNA gene sequence has also revealed that the strain PPA12 belongs to a species of the genus *Streptomyces*. The BLAST analysis showed 99% similarity to *Streptomyces fumigatiscleroticus* NRRL B-3865 deposited in Gen Bank with the accession number MK573175.

Findings of the present study showed that mangrove ecosystem of Parangipettai is a promising source for bioactive actinobacteria and the *Streptomyces fumigatiscleroticus* PPA12 from this ecosystem is a promising source for the isolation of bioactive compounds for clinical important carbapenam resistant *K. pneumoniae.*

ACKNOWLEDGEMENT:

Authors thank the authority of Sathyabama Institute of Science and Technology (Deemed to be University). Chennai, Tamil Nadu for the research facilities provided.

REFERENCES:

1. Ibrahim AH, Samar YD, Mostafa AF, Mohamed SK, Tobias AMG and Abdelmohsen UR. Natural Product Potential of the Genus *Nocardiopsis.* Marine Drugs. 2018; 16(5): 147.
2. Manivasagan P, Kyong-Hwa K, Sivakumar KECY, Li-Chan, Hyun-Myung Oh and Se-Kwon Kim. Marine actinobacteria: An important source of bioactive natural products. Environ. Toxicol. Pharmacol. 2014; 38(1):172-188.
3. Yu J, Liu Z, Qiu L, Xiaohui Q, Ying J and Beom SK. Isolation and characterization of actinobacteria from Yalujiang coastal wetland, North China. Asian Pacific J Tropical Biomed. 2015; 5(7):555-560.
4. Elbendary AA, Hessain AM, El-Hariri MD, Seida AA, Moussa IM, Mubarak AS, Kabli SA, Hemeg HA and El Jakee JK. Isolation of antimicrobial producing Actinobacteria from soil samples. Saudi J. Biol. Sci. 2018; 25(1):44-46.
5. Meletis G. Carbapenem resistance: overview of the problem and future perspectives. Therapeutic adv. Infectious Dis. 2016; 3(1):15-21.

6. Sathiyaseelan K and Stella D. Isolation, identification and antimicrobial activity of marine actinomycetes isolated from parangipettai. Rec. Res. Sci. Tech. 2011; 3(9): 74-77.
7. Radhakrishnan M, Balaji S and Balagurunathan R. Thermotolerant actinomycetes from the Himalayan Mountain-Antagonistic potential, characterization and identification of selected strains. Malaysian Appl. Bio. J. 2007; 36(1): 59-65.
8. Mohamed H, Miloud B, Zohra F, Garcia-Arenzana JM, Veloso A and Rodriguez-Couto S. Isolation and Characterization of actinobacteria from Algerian Sahara soils with antimicrobial activities. Int. J. Mol. Cellular Med. 2017; 6(2):109-120.
9. Radhakrishnan M, Pazhanimurugan R, Gopikrishnan V, Balagurunathan R and Vanaja Kumar. *Streptomyces sp* D 25 isolated from Thar Desert soil, Rajasthan producing pigmented antituberculosis compound only in solid culture. J. Pure Appl. Microbiol. 2014; 8(1): 333-337.
10. Sathish KS and Kokati VBR. *In-vitro* antimicrobial activity of marine actinobacteria against multidrug resistance *Staphylococcus aureus*. Asian Pacific J. Tropical Biomed. 2012; 2(10):787-792.
11. Eccleston GP, Brooks PR and Kurtboke DI. The occurrence of bioactive *Micromonosporae* in aquatic habitats of the sunshine coast in Australia. Marine Drug. 2008; 6243-6261.
12. Bavya M, Mohanapriya P, Pazhani murugan R and Balagurunathan R. Potential bioactive compounds from marine actinomycetes against biofouling bacteria. Indian J. Geomarine Sci. 2011; 40(4): 578-582.
13. Selvameenal L, Radakrishnan M and Balagurunathan R. Antibiotic pigment from desert soil actinomycetes; biological activity, purification and chemical screening. Indian J. Pharma Sci. 2009; 71: 499 - 504.
14. Augustine SK, Bhavsar SP and Kapadnis BP. A non-polyene antifungal antibiotic from *Streptomyces albidoflavus* PU23. J. Biosci. 2005; 30: 201 - 211.
15. Balagurunathan R and Radhakrishnan M. Biotechnological, genetic engineering and nanotechnological potential of actinomycetes. 2010; p. 436 In D.K. Maheshwari, R.C. Dubey, and R. Saravanamuthu (ed.), Industrial Exploitation of Microorganisms. I.K. International Publishing House Pvt. Ltd, New Delhi.
16. Shirling EB and Gottileb D. Methods for characterization of *Streptomyces* species. Int. J. Systematic Bacteriol. 1966; 6(3): 313 -340.
17. Gothwal RK, Nigam VK, Mohan MK, Sasmal D and Ghosh P. Screening of nitrogen fixers from rhizospheric bacterial isolates associated with important desert plants**.** Appl. Ecol. Environ Res. 2008; 6(2): 101-109.
18. Takahashi YO and Omura S. Isolation of new actinomycete strains for the screening of new bioactive compounds. J Gen Applied Microbiol. 2003; 49: 141-154.
19. Thompson JD, Higgins DG and Gibson TJ. Clustal W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22: 4673-4680.
20. Boudjella H, Bouti K, Zitouni A, Mathieu F, Lebrihi A and Sabaou N. Taxonomy and chemical characterization of antibiotics of *Streptosporangium* Sg 10 isolated from a Saharan soil. Microbiol. Res. 2006; 161: 288-298.
21. Felsenstein J. Phylogenies and the Comparative Method. The American Naturalist. 1985; 125(1): 1-15.
22. Berdy J. Thoughts and facts about antibiotics: where we are now and where we are heading. J. Antibiot. (Tokyo). 2012; 65(8): 385-395.
23. Tiwari, K. and Gupta, R.K. 2012. Rare actinomycetes: a potential storehouse for novel antibiotics. Crit. Rev. Biotechnol. 32(2):108-32.
24. Radhakrishnan M, Vijayalakshmi G, Gopikrishnan V and Jerrine J. Bioactive potential of actinobacteria isolated from certain understudied regions in India. J. Applied Pharma. Sci. 2016; 6(8): 151-155.
25. Mohana S and Radhakrishnan M. *Streptomyces* sp MA7 isolated from mangrove rhizosphere sediment effective against Gram negative bacterial pathogens. International Journal of PharmTech Research. 2014; 6(4): 1259-1264.
26. Diraviyam T, Radhakrishnan M and Balagurunathan R. Antioxidant activity of melanin pigment from Streptomyces species D5 isolated from Desert soil, Rajasthan, India Drug Invention Today. 2010; 3(3): 12-13.
27. Knight V, Sanglier JJ, DiTullio D, Braccili S, Bonner P, Waters J, Hughes D and Zhang L. Diversifying microbial natural products for drug discovery. Appl. Microbiol. Biotechnol. 2003; 62: 446-458.
28. Salamoni SP, Mann MB, Campos FS, Franco AC, Germani JC and Van der Sand ST. Preliminary characterization of some *Streptomyces* species isolated from a composting process and their antimicrobial potential. World J. Microbiol. Biotechnol. 2010; 26(10): 1847-1856.
29. Lalitha C, Thiagarajan R, Sudarshan S, Rathore and Jayapradha R. 2010. Bioactive molecule from *Streptomyces sp*. mitigates MDR *Klebsiella pneumoniae* in Zebrafish infection model. Front. Microbiol. 2010; 8: 614.
30. Treskatis SK, Orgeldinger V, Wolf H and Gilles ED. Morphological characterization of filamentous microorganisms in submerged cultures by on-line digital image analysis and pattern recognition. Biotechnology and Bioengineering. 1997; 53(2):191-201.
31. Manteca A, Alvarez R, Salazar N, Yague P and Sanchez J. Mycelium differentiation and antibiotic production in submerged cultures of *Streptomyces coelicolor.* Applied and Environmental Microbiology. 2008; 74(12): 3877-3886.

Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)
IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2

Android Live Text Recognition and Translation Application using Tesseract

Dr. S. Revathy
School of Computing
Sathyabama Institute of Science and Technology
Chennai, India
ramesh.revathy@gmail.com

Sakhayadeep Nath
Information Technology
Sathyabama Institute of Science and Technology
Chennai, India
sakhayadeepnath@gmail.com

ABSTRACT- There are thousands of languages in the world. Even though English is the most popular business language, only about twenty percent of the world population can speak English, moreover, people always prefer to use their language. Hence, translation is very important for advancements in literature, technology, and international politics. In this proposed work, we have made an application that can be used to take pictures of text in the paper, screen, signboards, or anywhere and the application will translate the text right there. The goal is to be able to understand any text in any language and translate it in real-time to any other language of preference. To achieve this, we have used Android as my preferred platform as it is a widely used mobile OS around the globe. We have also used many open-source projects to help us achieve our goal, such as Tesseract Leptonica for image computation, google-API-translate to translate the recognized text. We can recognize text from printed papers or screens and translate them into any language of preference. The accuracy in recognizing handwritten text is low and needs additional work. The proposed work has a scope to be used by visually impaired to recognize the text from anywhere and read the translated text to them.

Keywords *– Mobile; recognition; translation; language*

I. INTRODUCTION

Google translate can translate texts from any language to any other language, but the limitation is that we have to feed the text manually to the Google translate application. The proposed work overcomes this. There is also Microsoft translator but it supports far fewer languages and features than Google. Then there is iTranslate voice, which can translate voice-to-voice, but it supports even fewer languages and accuracy is not so good. In this proposed work, we have used different techniques, technologies from different sources, and combined them to form a solution to the above problems. The proposed system can recognize text from any sources automatically, and translate to various languages with precision.

Reference [11] the paper researched the exactness of Google Chinese-to-English interpretation from the points of view of custom and attachment with two correlations: Google interpretation with human expert interpretation, and Google interpretation with the Chinese source language. Results of Pearson correlations on formality and cohesion indicated Google English interpretation was exceptionally associated with both human English interpretation and the original Chinese texts.

In the proposed work, we have created a mobile application that will capture images of the text in any language and translate it into any language of the user's preference. By doing this recognition and translation this application can not only just translate the text but it also opens up a new world for the users. With this

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:20:53 UTC from IEEE Xplore. Restrictions apply.

application now users can visit new places anytime they want or start reading a new book of their choice in any language they want. The user can travel to a new location and just point their phone towards a sign on the street and voila, they have the translation on their screen.

The platform chosen for the development of this application is Android, the reason being android is the most widely used UNIX based Mobile Operating System with a wide variety of features, like low cost, user-friendly, easy to use, and its wide presence in the market. The IDE used for the development of this application is Android Studio along with Android SDK (Software Development Kit).

Libraries like Tesseract [12] is used for recognizing text from captured images.

II. RELATED WORKS

Harneet Singh and AnmolSachan in their paper [2] noted that the method distinguishes the text using the processing capability. Yet it is far from practical use hence, in the proposed paper we will implement a practical android application

"Lemao Liu, Atsushi Fujita, Masao Utiyama, Andrew Finch, and EiichiroSumita"in their paper [3] proposes an approach to train word-level TQE models using bilingual corpora, which are typically used in machine translation training and is relatively easier to access".

"Xujun Peng, Huaigu Cao, Krishna Subramanian, Rohit Prasad, and Prem Natarajan in their paper [4] proposed a novel automated image quality assessment method that predicts the degree of degradation on OCR"..
"Mande Shen and Hansheng Lei in their paper [5]presented a novel and cost-effective image preprocessing method".
"ChamilaLiyanage, ThiliniNadungodage, and RuvanWeerasinghe in their paper [6] described a set of training regimes for Tamil using the Tesseract engine that has enabled them to develop a robust Tamil OCR system. This describes the training regimes for the languages".

III. PROPOSED WORK

In this paper, we have proposed to create a mobile application that uses the principles of OCR and machine translation to recognize foreign language text and translate it into any language.

Methodology:

In this Architecture, the text object will be processed using different software and algorithms. The image would be captured using a phone camera and then the image will be processed using Leptonica [8] algorithms. After that, it will be passed to Tesseract [12] for recognition where the image will be converted to the relevant text it works using line segmentation algorithms, which works on the principle of locating the optimal succession of texts and gap areas within vertical zones by applying Viterbi path are identified using the Viterbi algorithm, the outcomes are the happenings observed in sequence. When the content lines are discovered, the baselines are fitted correctly utilizing a "quadratic spline".. The steps for recognizing texts are as follows:

Step 1: Splotches are prearranged keen on lines of text, the regions and the lines are examined for a static pitch and the related typescript
Step 2: the words are recognized and satisfactory word goes into distinguisher that is adaotive for training.

Step 3: on idnetifying the lines in the text the baselines are fitted.

Step 4: Splotches are partitioned to fit the baselines.
After the text is recognized, for translation it will be sent to the machine translation algorithms [9] of Google where it will be translated and sent back to the application where the translated text will be shown to the user.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:20:53 UTC from IEEE Xplore. Restrictions apply.



PROPOSED ALGORITHM

Step 1: Capture image using the phone camera, for input to the pre-processing algorithm.

Step 2: Background noise removal using threshold and morphology.

Step 3: Conversion of RGB image to grayscale by taking average values of the pixels for edge detection.

Step 4: The adaptive threshold to get the image threshold for different regions of the text image. Adapting to the lighting gradient.

Step 5: Image binarization to reduce the image to black and white to separate the background and text

Step 6: Text deskewing by skew determination of text images and correction.

Step 7: Image resizing to properly process the text data.

Step 8: Text extraction using the tesseract framework.

Step 9: For the translation of the extracted text, an API endpoint is called with the text parameters; the response is processed for the user.

Step 10: The translated text is shown to the user, on the output screen.

IV. RESULTS AND DISCUSSION

The different activities of the proposed system are shown in the below figures. These figures will help in understanding the flow of the system. The three main activities of the system are Capture activity, Reading the text, and Translation activity.

Capture Activity:

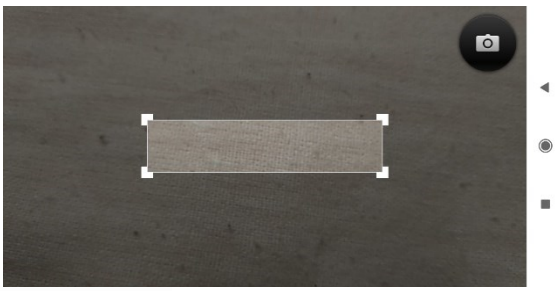


Fig 2: Capture activity

The first activity is the capture activity as shown in Fig 2; this opens the text recognizer, which is dependent on the camera of the android phone.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:20:53 UTC from IEEE Xplore. Restrictions apply.

It has a rectangular box, which can be adjusted to focus on the text we are interested in. This activity also has a capture button which when clicked captures the text within the rectangular box and translates the texts into the desired language.

Reading the Text:

In Fig 3, we can see that the desired text is being captured within the box. The box is adjustable according to the size or the length of the text we are interested in. This allows the user to capture any text they want without any limitations. This also helps in ignoring the unwanted texts from being captured by the application. This widget is very helpful for increasing the precision of the application.

Translation Activity:

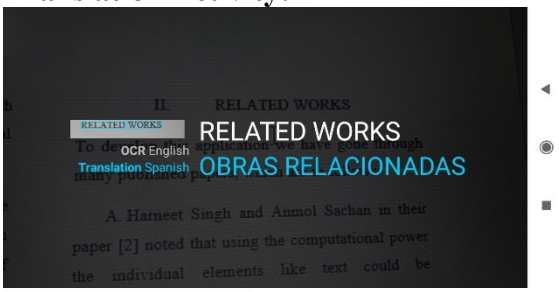


Fig 4: Translation activity

As shown in Fig 4, the next step is the translation of the recognized text. This is the ultimate goal of the application. This recognized text is sent for translation and the user receives the translated text for viewing. There is no limit to the size of the text that can be sent for translation. The device is used for capturing the image only limits this. The better the camera performance of the android device the more text can be accommodated in the viewing window of the application.

As seen in Fig 2, Fig 3 and Fig 4 the text is captured using the camera of the phone. It is processed to recognize text and then it is translated into different languages. Android software development kit is used to wrap the tesseract engine in an android mobile application to deliver a powerful yet simple and handy recognition and translation system to users.

In the proposed work, we have used various android compatible wrappers of Tesseract, leptonica, and Google's Android API for translation. Google services have been implemented. The image capture process is implemented using the inbuilt camera module in an android phone. This helps with the solution of real-time capture of images, and therefore no need for an already captured ready text image. Internet connectivity is essential as the proposed work relies on the internet for its working, i.e. the call made to the google API via the internet. "Android platform includes HttpsURLConnection client, which supports TLS, streaming uploads and downloads configurable timeouts, IPv6, and connection pooling. We have utilized the HTTP protocol to send and receive data".

The result is an android application that is capable of capturing the image in real-time via the camera module, extract the text in the image, and make an API call to google using the extracted text, and retrieve real-time translated text, which is shown to users.

V. PERFORMANCE OF THE PROPOSED SYSTEM

Fig 5 shows the performance of the proposed system. The proposed system was analyzed on the mean confidence of the OCR algorithm and the time taken by the system for the process. The average time required to recognize text and translate is less than 150 milliseconds. In this case, it is recognized with a mean confidence of 64.

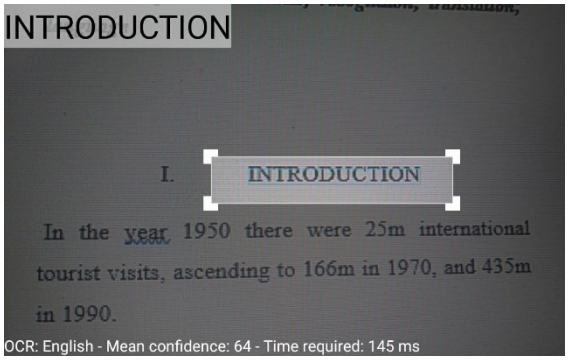


Fig 5: Performance of the proposed system

Table 1: Performance analysis of the proposed work

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:20:53 UTC from IEEE Xplore. Restrictions apply.

Parameters	Values
Mean confidence	> 60
Time required	< 150 ms
Other algorithms	>200 ms

As seen in Table 1, the mean confidence of text recognition is 64. The time required to recognize the same in less than150 ms, which is much less, than any other algorithms, which take much, more time for the process.

VI. CONCLUSION

Thus, this application is successfully able to recognize text from any language and then translate it into any other language any time anywhere. We believe this will be useful in many situations for users who want to discover new areas; it will be helpful for quick analysis of various texts in unknown languages.
This application uses various open-source libraries and algorithms for its purpose. This is used to successfully recognize text from various surfaces, backgrounds, and handwritten texts. The proposed architecture and the result is provided in the paper. In the future, this application can be extended to use for various other purposes such as for helping visually impaired to read the text [7], etc.

REFERENCES

[1]. Xiao Ling and Desmond Shaw, “Android Tesseract OCR. https://github.com/yushulx/android-tesseractocr
[2]. Singh, Harneet, and Anmol Sachan. "A Proposed Approach for Character Recognition Using Document Analysis with OCR." 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2018.
[3]. Liu, Lemao, et al. "Translation quality estimation using only bilingual corpora” IEEE/ACM Transactions on Audio, Speech, and Language Processing 25.9 (2017): 17621772.
[4]. Peng, Xujun, et al. "Automated image quality assessment for camera-captured OCR." 2011 18th IEEE International Conference on Image Processing. IEEE, 2011.
[5]. Shen, Mande, and Hansheng Lei. "Improving OCR performance with background image elimination." 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). IEEE, 2015.
[6]. Liyanage, Chamila, Thilini Nadungodage, and Ruvan Weerasinghe. "Developing a commercial-grade Tamil OCR for recognizing font and size-independent text." 2015 Fifteenth International Conference on Advances in ICT for Emerging Regions (ICTer). IEEE, 2015.
[7]. Manoharan, Samuel. "A Smart Image Processing Algorithm for Text Recognition Information Extraction and Vocalization for the Visually Challenged." *Journal of Innovative Image Processing (JIIP)* 1, no. 01 (2019): 31-38.
[8]. Leptonica.http://www.leptonica.org/
[9]. https://cloud.google.com/translate/
[10]. https://ourworldindata.org/tourism
[11]. Haiying Li, Arthur C. Graesser, Zhiqiang Cai, “Comparison of Google Translation with Human Translation”. Florida Artificial Intelligence Research Society Conference, 2014.
[12]. Tesseract Open Source OCR Engine
[13]. Kumar, Ritesh, Deb Jyothi, and S. Revathy. "Three Step Authentication Technique To Safeguard The Data Present In An Android Device." RESEARCH JOURNAL OF PHARMACEUTICAL BIOLOGICAL AND CHEMICAL SCIENCES 7, no. 4 (2016): 188193.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:20:53 UTC from IEEE Xplore. Restrictions apply.

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2020.1760136

Taylor & Francis
Taylor & Francis Group

Check for updates

Andrographolide as a potential inhibitor of SARS-CoV-2 main protease: an in silico approach

Sukanth Kumar Enmozhi[a], Kavitha Raja[a], Irudhayasamy Sebastine[b] and Jerrine Joseph[c]

[a]Department of Zoology, University of Madras, Chennai, India; [b]Department of Pharmacy, Periyar College of Pharmaceutical Sciences, Tiruchirappalli, India; [c]Centre for Drug Discovery and Development, Col. Dr. Jeppiaar Research Park, Sathyabama Institute of Science and Technology, Chennai, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

SARS-CoV-2 virus which caused the global pandemic the Coronavirus Disease- 2019 (COVID-2019) has infected about 1,203,959 patients and brought forth death rate about 64,788 among 206 countries as mentioned by WHO in the month of April 2020. The clinical trials are underway for Remdesivir, an investigational anti-viral drug from Gilead Sciences. Antimalarial drugs such as Chloroquine and Hydroxychloroquine derivatives are being used in emergency cases; however, they are not suitable for patients with conditions like diabetes, hypertension and cardiac issues. The lack of availability of approved treatment for this disease calls forth the scientific community to find novel compounds with the ability to treat it. This paper evaluates the compound Andrographolide from *Andrographis paniculata* as a potential inhibitor of the main protease of SARS-COV-2 (Mpro) through *in silico* studies such as molecular docking, target analysis, toxicity prediction and ADME prediction. Andrographolide was docked successfully in the binding site of SARS-CoV-2 Mpro. Computational approaches also predicts this molecule to have good solubility, pharmacodynamics property and target accuracy. This molecule also obeys Lipinski's rule, which makes it a promising compound to pursue further biochemical and cell based assays to explore its potential for use against COVID-19.

ARTICLE HISTORY
Received 8 April 2020
Accepted 20 April 2020

KEYWORDS
Andrographolide; in silico studies; SWISS-bioinformatics; SARS-CoV-2; plant compound

1. Introduction

The onset of symptoms such as fever, cough, fatigue, production of sputum, shortness of breath, sore throat, headache along with some with reports of diarrhoea and vomiting began to rise as the group of pneumonia cases from December 2019 and later they were identified as β-coronavirus in Wuhan, Hubei Province, China (Guan et al., 2020; Wang et al., 2020) The β- coronavirus was firstly named as 2019- novel coronavirus (2019-nCoV) on 12 January 2020 by WHO and formally named the disease as coronavirus 2019 (COVID- 19) and as a world emergency disease of cause and concern globally, International Committee of Coronavirus Study Group (CSG) recommended the use of the name as SARS- CoV- 2 which was published on 11 February 2020 (Guo et al., 2020). Through analyzing the viral sequence and evolutionary analysis, bat was suspected to be the natural host of the virus. And the virus might have been transferred to humans as their intermediate host by binding to ACE-2 Receptor (Angiotensin Converting Enzyme-2 Receptor) (Zhou et al., 2020). The first lethal case was reported on 11 January 2020. The infection from patients to healthcare workers was first verified on 20 January 2020. It was further reported that during Chinese New Year people migrated from Wuhan to various countries of the world. New cases evolved to other countries especially to various patients with no travel history to China which notified scientific and medical communities that local human to human transmission were seen in those countries (Rothe et al., 2020). Recently, the total number of cases around the world was recorded to be 1,203,959 confirmed cases with more than 64,788 deaths (https://www.worldometers.info/coronavirus/). Various types of treatments have been proposed which are mainly antiviral drugs, SARS- Cov and MERS- Cov antibodies are being used by clinicians and recent recommended combination therapy of hydroxychloroquine and azithromycin was studied and its results of open labelled non randomized clinical trial was published (Huang et al., 2020; Gautret et al., 2020).

Meanwhile, Food and Drug Administration (FDA) has stated that both Chloroquine phosphate and Hydroxychloroquine sulphate are not approved of treating COVID-19. And upon certain *in vitro* and some clinical data chloroquine phosphate and hydroxychloroquine sulphate was advised to be the treatment for COVID-19 and enough randomized trials on these compounds to be provided and allowed the administration of the above drugs to be used for emergency (https://www.fda.gov/emergency-use-authorization#covidtherapeutics). Hydroxychloroquine may have inhibitory mechanism over the viral processes and

CONTACT Jerrine Joseph jerrine.jj@gmail.com Centre for Drug Discovery and Development, Col. Dr. Jeppiaar Research Park, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu 600119, India.

Supplemental data for this article is available online at https://doi.org/10.1080/07391102.2020.1760136.

© 2020 Informa UK Limited, trading as Taylor & Francis Group

metabolisms. They may be involved in other mechanisms as inhibition of ACE2 cellular receptor, acidification of the cell membrane preventing the entry of virus and modulation of immune response through respective cytokine release (COVID-19 Drug Therapy-Elsevier, 09 March 2020). But recent studies have shown that the hydroxychloroquine can also cause drug poisoning and severe or moderate adverse effects in individuals who are already taking treatments for diabetic and hypersensitive patients, the same patient group who are found to be affected severely by COVID-19. Administration of hydroxychloroquine has found to inhibit pro-inflammatory cytokines which finally leads to Acute Respiratory Distress Syndrome (ARDS) (Guastalegname & Vallone, 2020). It has been found out that adverse neuropsychiatric condition was seen in post treatment of hydroxychloroquine which is hypothesized that it specifies the lysosomal dysfunction leading to psychiatric symptoms, which initiated the normal state of the patient who has been administered with the drug (Ali & Jones et al., 2018). Lethal adverse effect of retinal toxicity was seen in patient with acute renal impairment when administered with hydroxychloroquine (Tailor et al., 2012). A study of high doses of hydroxychloroquine along with atorvastatin in diabetic patients showed highest decline of blood glucose in patients (Wondafrash et al., 2020). When Antimalarial drug, hydroxychloroquine when administered to patients with dermatomyosis, non-life threatening cutaneous reactions are seen most in dermatomyosis patients than cutaneous lupus erythrematosus (Pelle & Callen, 2002) and many side effects has been reported. And according to the website, (https://www.guidetopharmacology.org/coronavirus.jsp) many ligands which are synthetic in nature have been proposed for the treatment of COVID-19 and are in clinical trials and are in process of peer review.

Due to these high adverse effects and the site of target through which Hydroxychloroquine acts on the viral proteasome, spike proteins and proteins involved in the life cycle of the virus are unknown (COVID-19 Drug Therapy-Elsevier, 09 March 2020). A potential natural, non- synthetic drug compound has to be found with minimal side effects. The drugs which are necessary to act on the targets such as ACE-2 receptors, TMPRSS2, SARS-CoV-2 and CD147 (https://www.guidetopharmacology.org/coronavirus.jsp) are in the process of being found in order to decrease the prognosis of the disease and life cycle of the virus. *In silico* studies of chemically synthetic drugs such as Paritaprevir and Raltegravir, Doultegravir and Bictegravir for the targets 3CLpro and 2′-OMTase (Khan, Jha, et al., 2020), theophylline and pyrimidone derivatives as possible inhibitors of RNA bound N terminal domain (Sarma et al., 2020) and Remdesivir, Saquinavir and Darunavir also with two natural compounds, flavone and coumarine derivatives for the inhibition of 3CL pro (Khan, Zia, et al., 2020) have been published. Though there are many targets are found for the treatment of COVID-19, the main protease (M^{pro}) of SARS- CoV-2 was chosen due to interest of treating infected patients, to stop the multiplication of virus within the cells, through which M^{pro} was involved in the release of polypeptides which are functional extensive proteolysis and cleavage of the enzyme itself from the sites of genome, pp1a and ppa1ab (Jin et al., 2020).

Plant compounds are an ideal of finding drug components of interest and most economical one to produce quickly as possible. This is known as the concept of repurposing the natural phytomolecules which will hasten the drug discovery process. During a search for such potent plant compounds we found a recent study on potential plant compounds which are able to inhibit the M^{pro} is in a process of publication (Khaerunnisa et al., 2020), while discussing the findings on the paper with group of Siddha and Ayurvedic doctors, they have advised us to examine the properties of the traditional available plant, several plant molecules obtained from medicinal plants were explored. The outcome was that *Andrographis paniculata* which was found to have anti-viral property and already reported in ancient texts could be investigated as a good bet. We examined the main compound in *Andrographis paniculata* and found that the main compound in the plant was Andrographolide. The plant compound has been evidenced to have anti-inflammatory, anti-cancer, anti-obesity and anti-diabetes (Dai et al., 2019). Andrographolide was found to have antiviral properties over many types of viral infections (Gupta et al., 2017) and was found to have activity against Chikungunya (Wintachai et al., 2015) potential inhibitor of herpes simplex virus type-1 (Seubsasana et al., 2011). Also, during the outbreak of dengue in India at 2006, aqueous extracts of *Andgrographis paniculata* were given through the advice of Ministry of AYUSH (Ayurveda, Unani, Siddha and Homeopathy) department of India which led to decrease in cases and infection of the disease as a preventive measure even to normal people acting as a immune booster. The study of anti-dengue activity was found and published through quantification of dengue viral inhibition and showed most antiviral inhibitory effects when DENV 1-4 infected Vero cells with maximum non-toxic dose (Krishnasamy et al., 2018). Due to these antiviral activities the compound Andrographolide was chosen.

Due to these interests, this paper involves the *in silico* analysis of Andrographolide against crystal structure of SARS-CoV-2 main protease which is provided with an inhibition site (PDB ID: 6LU7) (Zhang et al., 2020), prediction of ADME (https://www.swissadme.ch), Target Prediction (https://www.swisstargetprediction.ch) were done by using Swiss-Bioinformatics online Tools. The prediction of toxicity of Andrographolide was seen using pkCSM online web tool (https://biosig.unimelb.edu.au/pkcsm/).

2. Materials and methods

2.1. Docking of andrographolide on sars-cov-2 main protease

2.1.1. Receptor preparation

The SARS-CoV-2 main protease (Figure 1a and 1b) (PDB ID: 6LU7) (Jin et al., 2020) was used as the receptor. The inhibitor was selected and removed. The receptor preparation was done using the Dock Prep tool of UCSF-Chimera. Hydrogens were added and optimized by a hydrogen bonding network and allowed the method to determine the Histidine protonation state. The receptor was saved in mol2 format

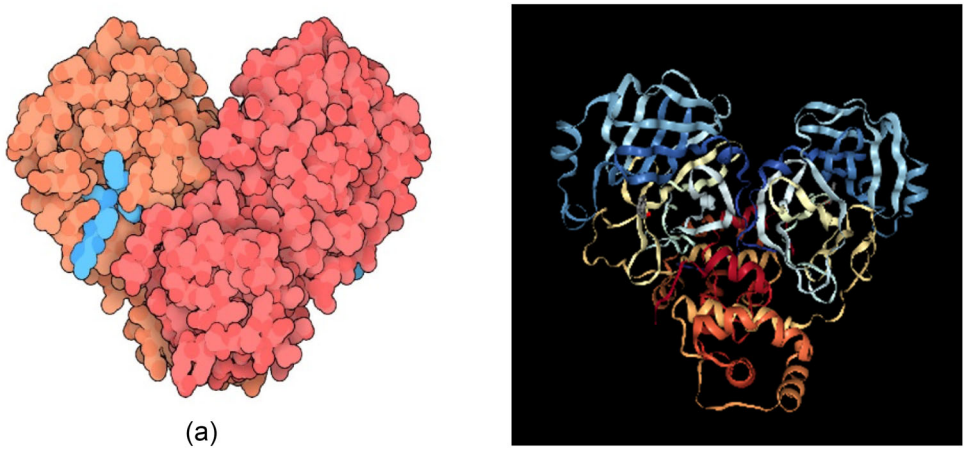


Figure 1. (a) SARS- CoV-2 main protease with inhibitor in turquoise; (b) Crystal structure of SARS- CoV-2 main protease with inhibitor.

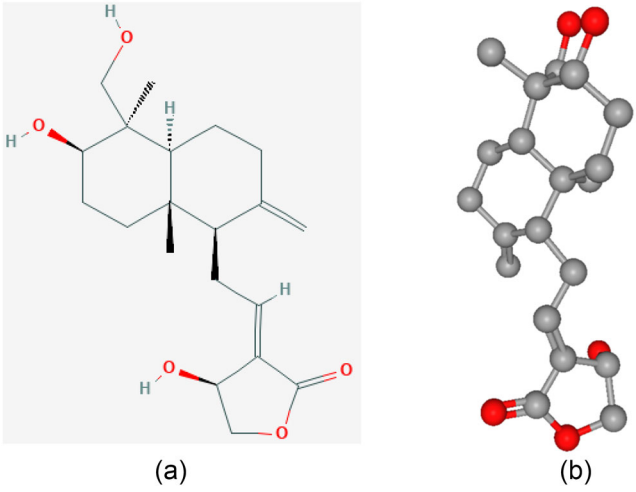


Figure 2. (a) 2D structure of andrographolide; (b) 3D structure of andrographolide.

(rec_charged.mol2) using untransformed coordinates and Sybyl-style hydrogen naming. Hydrogen present in the structure were removed from the protein and saved in pdb format (rec_noH.pdb).

2.1.2. Ligand preparation

The 3 D structure of Andrographolide (Figure 2a and 2b) (PubChem CID:5318517) was downloaded from Pubchem (https://pubchem.ncbi.nlm.nih.gov/compound/Andrographolide). Hydrogen was added to the ligands using the AddH function in structure editing tools in UCSF-Chimera. Charges were added to the ligands using AM1-BCC method. The ligand file was saved in mol2 format (lig_charged.mol2) using untransformed coordinates.

2.1.3. Docking

The molecular surface of the receptor (rec_noH.pdb) was prepared using the Write DMS Tool in UCSF-Chimera. Spheres outside the surface were generated using sphgen function having 4.0 in Angstroms as the maximum sphere radius and 1.4 in Angstroms as the minimum sphere radius. The cluster present in the binding site of the inhibitor was chosen using showsphere function. A box was created around the chosen cluster, having extra margins enclosed of 5.0 in Angstroms in all the 6 directions. A grid was generated using the program grid of DOCK6. Flexible docking parameters were employed for Andrographolide.

2.2. Adme prediction

ADME (Adsorption, Distribution, Metabolism and Excreation) is important to analyze the pharmacodynamics of the proposed molecule which could be used as a drug. SWISS-ADME is a website (https://www.swissadme.ch) which allows the user to draw their respective ligand or drug molecule or include SMILES data from PubChem and provides the parameters such as lipophilicity (iLOGP, XLOGP3, WLOGP, MLOGP, SILICOS-IT, Log P_0/w), water solubility- Log S (ESOL, Ali, SILICOS-IT), drug likeness rules (Lipinski, Ghose, Veber, Egan and Muegge) and Medicinal Chemistry (PAINS, Brenk, Lead-likeness, Synthetic accessibility) methods are analyzed (Daina et al., 2017). The data from PubChem which consists of SMILES of Andrographolide (https://pubchem.ncbi.nlm.nih.gov/compound/Andrographolide) was entered into the search bar and was analyzed.

2.3. Target prediction

Molecular Target studies are important to find the phenotypical side effects or potential cross reactivity caused by the action of small biomolecules (Keiser et al., 2007; Gfeller et al., 2014). Swiss Target Prediction website (https://www.swisstargetprediction.ch) was logged on and the ZINC number for Andrographolide (ZINC3881797) was entered on to the search bar and was analyzed.

2.4. Toxicity prediction

Toxicology prediction of small molecules is important to predict amount of tolerability of the small molecule before

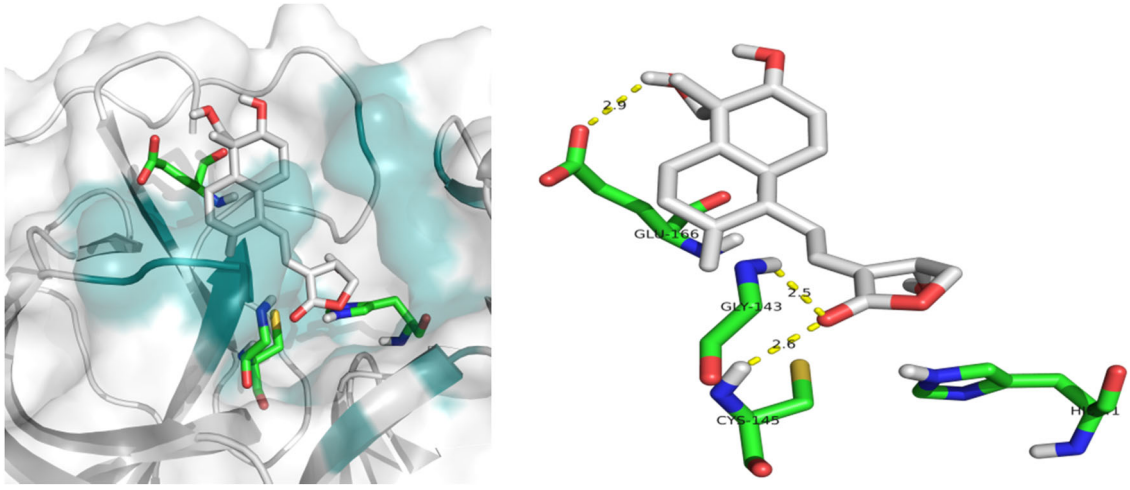


Figure 3. (a) Docked Ligand in the binding pocket of the protein. The binding pocket of the protein is in light blue and the surface of the protein in white; (b) Residues interacting with the ligand and their hydrogen bonds distance. The Ligand is given in white, interacting residues in green and hydrogen bonds in yellow.

TABLE 1. Interaction of the SARS-CoV2 Main Protease with Andrographolide Ligand.*

Compound	No of H- Bonds	Residue Receptor	Ligand	Bond Length (Å)	Docking Score (Kcal/ mol)
Andrographolide	**4**	**Cys145(H)**	**O2**	**2.46**	**-3.094357**
		Gly143(H)	**O2**	**2.62**	**-3.094357**
		Glu166(H)	**O3**	**2.93**	**-3.094357**
		Glu166(H)	**O3**	**2.92**	**-3.094357**

*Prediction of noncovalent interactions for PDB structure using PLIP v1.4.4 (Salentin,S. et al. PLIP: fully automated protein-ligand interaction profiler. Nucl. Acids Res. (1 July 2015) 43 (W1): W443-W447. doi: 10.1093/nar/gkv315).

being ingested into the human and animal models. pkCSM is an online database in which the small molecule can be drawn virtually or can be analyzed by submitting the SMILES of the same. The website can provide details of toxicology effects in the fields of AMES Toxicity, human maximum tolerance dose, hERG-I inhibitor, hERG-II inhibitor, LD50, LOAEL, Hepatotoxicity, Skin Toxicity, *T. pyriformis* toxicity and Minnow toxicity. The website was logged on and the SMILES of the Andrographolide data from PubChem was searched and submitted into the website and toxicity mode was selected (Pires et al., 2015).

3. Results

3.1. Docking

The docking analysis of the compound with SARS-CoV-2 protease generated negative values for free energy -3.094357 KJ/mol in the grid box, suggesting high affinity for the binding pocket. All the binding conformations of the compound in the active binding pocket involved both H-bond and salt bridge interaction. The compound did bind to the protease with 4 hydrogen bonds with 3 residues namely Gly143, Cys145 and Glu166 as shown (Figure 3a and 3b). All details of the atoms involved in bonding with ligands, bond lengths, docking energies and salt bridges are given in Table 1.

3.2. Adme prediction

The ADME prediction which was done using SWISSADME database came with the results following after submission of the small biomolecule, Andrographolide. The physiochemical properties of the compound are, number of 25 heavy atoms, 5 hydrogen bond acceptors, 3 hydrogen bond donors, molar refractivity of 95.21and topological polar surface area (TPSA) of the molecule is found to be 86.99Å^2. The lipophilicity of the molecule, iLOGP is 2.45, XLOGP3 is 2.16, WLOGP is 1.96, MLOGP is 1.98, SILICOS-IT is 2.94 and Consensus P_0/W is 2.30.

Water Solubility properties calculated are ESOL -3.18, solubility of 2.36e^{-02}mg/ml and of soluble class; Ali -3.62, solubility of 8.42e^{-02}mg/ml and of soluble class, SILICOS-IT -2.69, solubility of 7.22e^{-01}mg/ml and of soluble class. Pharmacokinetic data predicted was found to be of high Gastrointestinal absorption (GI), not blood brain barrier permeant, acts as a P-gp substrate, does not inhibit CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4 cytochromes. Skin permeation kinetics (Log K_p) was found to be -6.90 cm/s.

Druglikeness factors was found to be of drug like compound which obeys Lipinski's Rules with no violation, also obeys druglikeness score rules such as Ghose, Veber, Egan, Muegge and with 0.55 Bioavailability score. Medicinal chemistry parameters were found to be of no PAINS alert, violates Brenk's laws with two alerts of being an isolated alkene, one michael acceptor, no lead likeness with molecular weight of greater than 350 and synthetic accessibility of 5.06 rate.

3.3. Target prediction

The target prediction analysis was displayed in the Web page with the following observations the top 25 of the

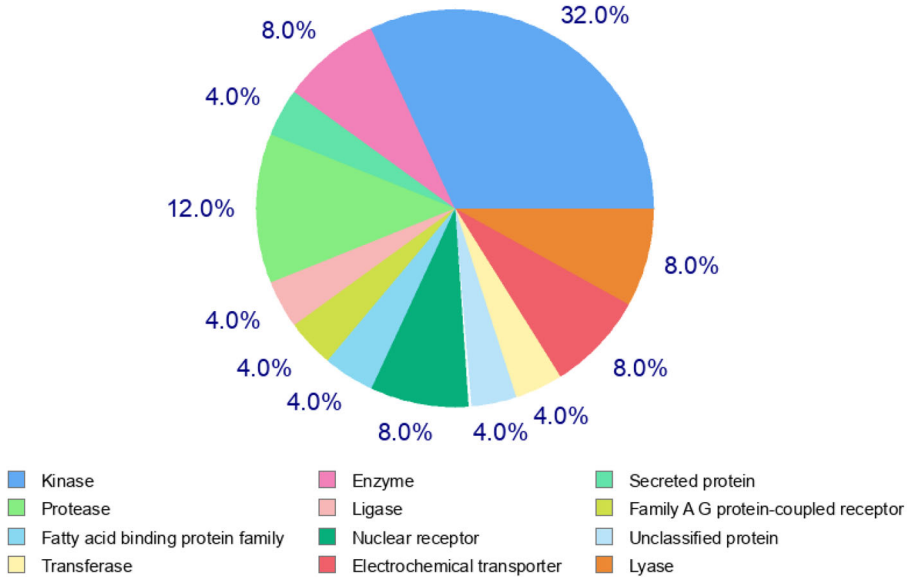


Figure 4. Top-25 of Target Predicted for Andrographolide.

results were given as a pie-chart (Figure 4). The pie chart predicts 32% of Kinase, 12% of protease, 4% of Fatty acid binding protein family, 4% of Transferases, 8% of Enzymes, 4% of Ligases, 8% of Nuclear Receptors, 8% of electrochemical transporters, 4% of Secreated protein, 4% of Family AG protein coupled receptor, 4% of unclassified protein and 8% of Lyase. The output table consistting of Target, Common Name, Uniprot ID, ChEMBL-ID, Target Class, Probability and Known actives in 2 D/3D are given in the Supplementary-1. The possible sites of target which the compound may bind to are mostly the targets which are predicted by the software and the probability score are very less that is from 0.10560 to 0.0972. This makes an inference that the small compound may have high target attraction towards the specific binding site it is directed to.

3.4. Toxicity prediction

The toxicity predicted was displayed in the website and the results is as follows, The Andrographolide does not have AMES toxicity, Maximum tolerated dose for human is about 0.128 log mg/kg/day, it does not inhibit hERG-I and hERG-II, Acute oral rat toxicity (LD_{50}) was found to be 2.162 mol/kg, Chronic oral rat toxicity (LOAEL) was found to be 1 log mg/kg_bw/day, does not produce hepatotoxicity, it does not cause skin sensitivity, 0.491 log μg/L causes *T. pyriformis* toxicity and 1.37 log mM causes Minnow toxicity.

4. Discussion

The need of the hour is a therapy for SARS-CoV-2 virus, many small molecules like Remdesivir are in trial to provide cure for this dreadful viral outbreak. Hydroxychloroquine and azithromycin complex is being advised to be given for the affected in case of emergency, though it has been studied to increase the pH of the protease and is published as a potent inhibitor of SARS-CoV-2 infection and spread (Wang et al., 2020; Heald-Sargent & Gallagher, 2012; Vincent et al., 2005). The Hydroxychloroquine and azithromycin complex mentioned above may be potent, but the adverse side effects they bring to the patients are very alarming as shared already, with these drastic side effects, the need to bring equally or more potent alternatives in the form of plant derived drug is very essential as they are much safe and have no known side effects. As we are interested only in finding pure potent plant compounds without adding any analogs or derivatives we found that the plant compound, Andrographolide intriguing due to its awesome properties. The drug compound, Andrographolide can be isolated and produced easily by extracting from the plant *Andrographis paniculata*. When the compound was analyzed by *in silico* computational docking tools it successfully docked against the inhibitor region of the main protease of SARS-CoV-2 virus with docking score of -3.094357 Kcal/mol, the docking score showed great binding when compared to synthetic compounds when they are docked against M^{pro} such as disulfiram, tideglusib and shikonin which are -46.16 Kcal/mol, −61.79 Kcal/mol and -17.35 Kcal/mol (Jin et al., 2020). And it also shows great binding score when compared against recently proposed combination of three drugs namely, lopinavir, ostelmivir and ritonavir whose binding scores are -4.1Kcal/mol,-4.65 Kcal/mol and -5.11 Kcal/mol (Muralidharan et al., 2020). Even some plant molecules which are studied to inhibit the main protease of SARS-CoV-2 failed to prove their binding score when compared with the Andrographolide. They are compounds such as kaempferol -9.41 Kcal/mol, quercetin -8.58 Kcal/mol, demethoxycurcumin -8.17 Kcal/mol, curcumin -7.31 Kcal/mol, catechin –7.05 Kcal/mol, epichatechin gallate- 7.24 Kcal/mol, zingerol -6.67 Kcal/mol and gingerol -5.40 Kcal/mol respectively (Khaerunnisa et al., 2020). Even proposed inhibitor of M^{pro} such as PRD_002214 has a docking score of -10.466 Kcal/mol (Bouchentof & Missoum, 2020) which proclaims that Andrographolide has better properties than other proposed inhibitors.

The compound possesses excellent properties of druggability as well small biomolecule. The molar refractivity of the compound confirms that the drug compound is permeable through particular membranes and can remain constant even in the midst of strong or weak solute-solvent, solvent-solvent interactions. And exemplary TPSA says that it has great transport properties. Through lipophilicity of the drug compound we can know that the compound has ideal property for oral and intestinal absorption and is able to be absorbed sub-lingual as well. Through water solubility properties predicted the drug is free soluble.

The pharmacokinetic data predicted about absorption and permeability echoes the predicted values of lipophilicity and solubility, they also predict that is able to release phosphate from ATP and simultaneous binding of ADP to the glycoprotein, thus the compound acts as a p-gp substrate. The compound, Andrographolide does not inhibit liver metabolism by inhibiting CYP1A2, it does not stop the metabolism of several therapeutic drugs especially, anti-ulcer, anti- malarial, anti-convulsant, anesthetic and sedative drugs via inhibiting CYP2C19, it does not stop metabolism of anti-hypersensitive drugs, β blockers, anti-arrhythmic drugs and anti-depressants via inhibition of CYP2D6, it does not stop the metabolism of anti-clotting agents, anti-seizure, management of type-II diabetes, anti-hypertensive, Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) via CYP2D9 and it does not stop oxidation of steroids, fatty acids and xenobiotics as well as for hormone synthesis and breakdown through CYP3A4.

Druglikeness factor rules were obeyed accordingly without any violation to this compound which describes the compound can act as a drug in the biological systems. Medical Chemistry parameters exclaims that, zero Pan-Assay Interference Compounds (PAINS) alert for the compound meaning that it is a progressive compound worthy of testing for biochemical assays. The toxicity prediction says that the compound, Andrographolide is safe and can be given as a drug with the value of tolerance prescribed for human consumption as predicted by the website. Though these properties are appreciable *in silico*, due to extensive lockdown and work from home command through the government the findings was not continued in a wet lab. So, further studies of *in vitro* and clinical studies dealing with SARS-CoV-2 should be considered for further studies.

Acknowledgements

I would like to express my sincere gratitude to Derek and Sharon for their daily updates on SARS-CoV-2, Dr. Kaleeswaran for his advice on this compound and Mr. Rahul Vivek for his work on docking.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Ali, S. S., & Jones, H. (2018). 23. An adverse neuropsychiatric reaction following treatment with hydroxychloroquine: A case report. *Rheumatology Advances in Practice*, *2*(suppl_1). https://doi.org/10.1093/rap/rky033.014

Bouchentouf, S., & Missoum, N. (2020). Identification of compounds from Nigella Sativa as new potential inhibitors of 2019 novel Coronasvirus (COVID-19): Molecular docking study. https://doi.org/10.26434/chemrxiv.12055716

Dai, Y., Chen, S., Chai, L., Zhao, J., Wang, Y., & Wang, Y. (2019). Overview of pharmacological activities of Andrographis paniculata and its major compound andrographolide. *Critical Reviews in Food Science and Nutrition*, *59*(sup1), S17–S29. https://doi.org/10.1080/10408398.2018.1501657

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*, 42717. https://doi.org/10.1038/srep42717

Gautret, P., Lagier, J., Parola, P., Hoang, V. T., Meddeb, L., Mailhe, M., Doudier, B., Courjon, J., Giordanengo, V., Vieira, V. E., Dupont, H. T., Honoré, S., Colson, P., Chabrière, E., La Scola, B., Rolain, J., Brouqui, P., & Raoult, D. (2020). Hydroxychloroquine and azithromycin as a treatment of COVID-19: Results of an open-label non-randomized clinical trial. *International Journal of Antimicrobial Agents*, 105949. https://doi.org/10.1016/j.ijantimicag.2020.105949

Gfeller, D., Grosdidier, A., Wirth, M., Daina, A., Michielin, O., & Zoete, V. (2014). SwissTargetPrediction: A web server for target prediction of bioactive small molecules. *Nucleic Acids Research*, *42*(W1), W32–W38. https://doi.org/10.1093/nar/gku293

Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., Liu, L., Shan, H., Lei, C. L., Hui, D., Du, B., Li, L. J., Zeng, G., Yuen, K. Y., Chen, R. C., Tang, C. L., Wang, T., Chen, P. Y., Xiang, J., Li, S. Y, … China Medical Treatment Expert Group for Covid-19 (2020). Clinical characteristics of Coronavirus disease 2019 in China. *The New England Journal of Medicine*, NEJMoa2002032. Advance online publication. https://doi.org/10.1056/NEJMoa2002032

Guastalegname, M., & Vallone, A. (2020). Could chloroquine/Hydroxychloroquine be harmful in coronavirus disease 2019 (COVID-19) treatment? *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Sosciety of America*, ciaa321. Advance online publication. https://doi.org/10.1093/cid/ciaa321

Guo, Y., Cao, Q., Hong, Z., Tan, Y., Chen, S., Jin, H., Tan, K., Wang, D., & Yan, Y. (2020). The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. *Military Medical Research*, *7*(1),11. https://doi.org/10.1186/s40779-020-00240-0

Gupta, S., Mishra, K. P., & Ganju, L. (2017). Broad-spectrum antiviral properties of andrographolide. *Archives of Virology*, *162*(3), 611–623. https://doi.org/10.1007/s00705-016-3166-3

Heald-Sargent, T., & Gallagher, T. (2012). Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. *Viruses*, *4*(4), 557–580. https://doi.org/10.3390/v4040557

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., … Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan. *The* Lancet, *395*(10223), 497–506. https://doi.org/10.1016/s0140-6736(20)30183-5 https://doi.org/10.1016/S0140-6736(20)30183-5

Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L., Peng, C., Duan, Y., Yu, J., Wang, L., Yang, K., Liu, F., Jiang, R., Yang, X., You, T., Liu, X., … Yang, H. (2020). Structure of mpro from COVID-19 virus and discovery of its inhibitors. *Nature*. https://doi.org/10.1038/s41586-020-2223-y

Keiser, M. J., Roth, B. L., Armbruster, B. N., Ernsberger, P., Irwin, J. J., & Shoichet, B. K. (2007). Relating protein pharmacology by ligand chemistry. *Nature Biotechnology*, *25*(2), 197–206. https://doi.org/10.1038/nbt1284

Khaerunnisa, S., Kurniawan, H., Awaluddin, R., Suhartati, S., Soetjipto, S. (2020). Potential inhibitor of COVID-19 main protease (Mpro) from several medicinal plant compounds by molecular docking study. *Preprints*, 2020030226. https://doi.org/10.20944/preprints202003.0226.v1

Khan, R. J., Jha, R. K., Amera, G., Jain, M., Singh, E., Pathak, A., Singh, R. P., Muthukumaran, J., & Singh, A. K. (2020). Targeting SARS-Cov-2: A systematic drug repurposing approach to identify promising inhibitors against 3C-like Proteinase and 2′-O-RiboseMethyltransferase. *Journal of Biomolecular Structure and Dynamics*, 1–14. https://doi.org/10.1080/07391102.2020.1753577

Khan, S. A., Zia, K., Ashraf, S., Uddin, R., & Ul-Haq, Z. (2020). Identification of chymotrypsin-like protease inhibitors of SARS-Cov-2 via integrated computational approach. *Journal of Biomolecular Structure and Dynamics*, 1–10. https://doi.org/10.1080/07391102.2020.1751298

Krishnasamy, K., Ramalingam, S., Karupannan, S., Padmanaban, P., Vijayan, S., Sheriff, K., & Palani, G. (2018). Anti-dengue activity of Andrographis paniculata extracts and quantification of dengue viral inhibition by SYBR green reverse transcription polymerase chain reaction. *AYU (an International Quarterly Journal of Research in Ayurveda))*, *39*(2), 87. https://doi.org/10.4103/ayu.AYU_144_17

Muralidharan, N., Sakthivel, R., Velmurugan, D., & Gromiha, M. M. (2020). Computational studies of drug repurposing and synergism of lopinavir, oseltamivir and ritonavir binding with SARS-Cov-2 protease against COVID-19. *Journal of Biomolecular Structure and Dynamics*, 1–6. https://doi.org/10.1080/07391102.2020.1752802

Pelle, M. T., & Callen, J. P. (2002). Adverse cutaneous reactions to Hydroxychloroquine are more common in patients with Dermatomyositis than in patients with cutaneous lupus Erythematosus. *Archives of Dermatology*, *138*(9), 1231–1233. https://doi.org/10.1001/archderm.138.9.1231

Pires, D. E., Blundell, T. L., & Ascher, D. B. (2015). PkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, *58*(9), 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Froeschl, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., Seilmaier, M., Drosten, C., Vollmar, P., Zwirglmaier, K., Zange, S., Wölfel, R., & Hoelscher, M. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, *382*(10), 970–971. https://doi.org/10.1056/NEJMc2001468

Sarma, P., Sekhar, N., Prajapat, M., Avti, P., Kaur, H., Kumar, S., Singh, S., Kumar, H., Prakash, A., Dhibar, D. P., & Medhi, B. (2020). In-silico homology assisted identification of inhibitor of RNA binding against 2019-nCoV N-protein (N terminal domain). *Journal of Biomolecular Structure and Dynamics*, 1–11. https://doi.org/10.1080/07391102.2020.1753580

Seubsasana, S., Pientong, C., Ekalaksananan, T., Thongchai, S., & Aromdee, C. (2011). A potential Andrographolide analogue against the replication of herpes simplex virus type 1 in Vero cells. *Medicinal Chemistry*, *7*(3), 237–244. https://doi.org/10.2174/157340611795564268

Tailor, R., Elaraoud, I., Good, P., Hope-Ross, M., & Scott, R. A. (2012). A case of severe hydroxychloroquine-induced retinal toxicity in a patient with recent onset of renal impairment: A review of the literature on the use of Hydroxychloroquine in renal impairment. *Case Reports in Ophthalmological Medicine*, *2012*, 1–3. https://doi.org/10.1155/2012/182747

Vincent, M. J., Bergeron, E., Benjannet, S., Erickson, B. R., Rollin, P. E., Ksiazek, T. G., Seidah, N. G., & Nichol, S. T. (2005). Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virology Journal*, *2*(1), 69.https://doi.org/

Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., & Xiao, G. (2020). Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. *Cell Research*, *30*(3), 269–271. https://doi.org/10.1038/s41422-020-0282-0

Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., Zhao, Y., Li, Y., Wang, X., & Peng, Z. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in Wuhan, China. *JAMA*, *323*(11), 1061. https://doi.org/10.1001/jama.2020.1585

Wintachai, P., Kaur, P., Lee, R. C., Ramphan, S., Kuadkitkan, A., Wikan, N., Ubol, S., Roytrakul, S., Chu, J. J., & Smith, D. R. (2015). Activity of andrographolide against chikungunya virus infection. *Scientific Reports*, *5*(1), 14179. https://doi.org/10.1038/srep14179

Wondafrash, D. Z., Desalegn, T. Z., Yimer, E. M., Tsige, A. G., Adamu, B. A., & Zewdie, K. A. (2020). Potential effect of Hydroxychloroquine in diabetes mellitus: A systematic review on preclinical and clinical trial studies. *Journal of Diabetes Research*, *2020*, 1–10. https://doi.org/10.1155/2020/5214751

Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K., & Hilgenfeld, R. (2020). Crystal structure of SARS-Cov-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science (New York, N.Y)*, *368*(6489), 409–412. https://doi.org/10.1126/science.abb3405

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., … Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Analysis of the Water Quality Monitoring System

L. Lakshmanan, Jesudoss A, Sivasangari A, Sardar Maran and Mercy Theresa M

***Abstract*—Now a days many people are suffering from dangerous diseases which are caused due to impure water. In our project we are doing analysis for water quality monitoring system, it gives data about the quality of water, on a webpage. The quality of water is determined using various sensors like PH sensor and turbidity sensor, connected to the Arduino family microcontroller. The Arduino software is written in embedded C and GSM module is connected to the Arduino. The data will be transferred constantly from the remote sensor organize through microcontroller and wifi. Wifi module is used to send data to the webpage via internet which is connected to the microcontroller. The total data regarding the purity of water is displayed in the webpage and is analysed in the form of graph, pie chart and values are given in the table. We transfer this information to cloud and clients can get to this information through web page application, client from anyplace can screen the data whenever.**

***Index Terms*—Microcontroller, PH sensor, turbidity sensor, GSM and cloud storage.**

I. Introduction

In the present days and current period, we are moving towards making our urban communities as the brilliant urban areas, because many innovative research and developments throughout the decades. So the present period is said to be time of creations, time of improvement, time of globalization and the time of astuteness and so on. In any case, the counter side of the equivalent is that the present time is time of the contamination, a dangerous atmospheric deviation, weakness and hopeless wellbeing factors. One of the inborn and prime hindrance is total population does not have purified and safe water for drinking. This is increasing unsafe circumstances in some nations like India, where grimy water is being utilized for drinking with no appropriate water treatment before drinking. The fundamental driver for this are the numbness of individuals and government area and the inadequate water quality checking framework, which results in genuine medical problems.

Lakshman is a Professor, Dept. of CSE, Sathyabama Institute Of Science and Technology, Chennai (author email:lakshmanan.cse@sathyabama.ac.in)

Jesudoss A., Sivasangari A., and Sardar Maran are Associate Professors, Dept. of CSE, Sathyabama Institute Of Science and Technology, Chennai (corresponding author e-mail: jesudossas@gmail.com co-authors email: sivasangarikavya@gmail.com, sardarmaran@gmail.com).

Mercy Theresa is an Associate Professor, Dept. of ECE, Prince Shri Venkateshwara Padmavathy Engineering College, Chennai (author email:mercyjesudossa@gmail.com).

The inspiration of the proposed framework was to plan a remote framework to screen water quality in a most straightforward and practical way. This framework can break down some essential variables of water to take preventive measures for water quality support. The pH sensor and turbidity sensor are utilized to gather the pH and turbidity dimension of the water. With the utilization of wifi module, we can get the information from the rural and less developed areas. The sensors have the simple yield, consequently they are interfaced to simple contribution of the Nodemcu microcontroller and the information are exchanged through the wifi module. The PH and turbidity parameters thus calculated are stored in a tabular format and is shown on pc. The parameters that are utilized to decide the nature of the water are the pH level and turbidity level.

The rest of the paper is organized as follows. The related works are explained in Section II. Section III, Section IV and Section V describes about the system architecture, software and hardware specification respectively. The system design and analysis are discussed in Section VI and VII respectively. At last, Section VIII concludes the paper with conclusion of the work.

II. Related Works

Said M.F. et al describes the possibility of submerged remote sensor organize is the water quality monitoring utilizing remote sensor arrange innovation controlled by sunlight based board [1]. The hubs and the base stations are associated utilizing WSN innovation like Zigbee Data assembled by various sensors at the center point side, for instance, pH, Information gathered from the remote site can be shown in visual configuration a well as it very well may be examination utilizing distinctive reenactment instruments at base station. Lakshmanan et al indicates the significance of IoT and their benefits and also threats handled by IoT [2].

Ayushi S jaiswal, vaidehi Baporikar The data which is very difficult for humans to gather can be done by underwater robots. They are utilized broadly by mainstream researchers to think about sea submerged condition. Zigbee is a productive and viable remote system standard for remote control and checking applications. They displayed a reasonable and productive model of embedded remote information resource framework utilizing Zigbee which will be constrained by the PIC microcontroller [3].

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

IEEE Advancing Technology for Humanity

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:16:13 UTC from IEEE Xplore. Restrictions apply.

F Ntambi,CP kruger,BJ silva,GP Hancke,This idea implements the structure of a water the board framework that can log pH and weight readings remotely. This likewise decides a water spill in the pipe framework or water is sufficiently sheltered or human utilization.

A graphical UI was additionally actualized to show the information got and alert if any anomalies happened [4]. Jesudoss et al show how different types of sensors can be effectively implemented in safe driving [5].

Ji Wangi proposed A remote sensor systems for water quality checking. In which a novel arrangement of remote water quality estimating and observing dependent on remote sensor system and CDMA innovation [6].

Cheng-liang lai proposed utilizing picture preparing innovation for water quality checking framework. In which the effectively fabricated a water quality checking framework by using the picture handling framework and fluffy induction in auto perceiving the motion of fish [7]. The work done by Lakshmanan et al highlights the possible innovative communication methodologies and technologies for IoT [8]. Aravindan et al proposes the water quality management for real-time data [9]. Doukkali et al proposes MAC protocol using underwater acoustic sensor [10]. Akyildiz et al proposes underwater sensor nodes for autonomous underwater vehicles [11]. Jesudoss et al describes how IoT sensors are used in designing a smart helmet for accident avoidance [12]. Vasilescu et al also used underwater sensor networks for tracking coral reefs and fisheries [13] . Niel Andre Cleote et al explains the techniques for designing smart sensors for checking quality of the water [14]. Jiang P. et al proposes another efficient design for checking the water quality using wireless sensor networks [15]. Fisher et al provides a complete insight about the wireless sensor networks and its effective implementation [16]. Lambrro et al experiments the water quality for nephelometric turbidity systems and provides a different approach [17]. Vinod Raut et al uses an innovate wireless technique for monitoring the water quality [18]. Yi Hong Wang is another motivational work done for monitoring the quality of the water [19].

III. System Architecture

The micro controller block diagram is shown in Fig. 2. Probably every individual is aware that water is one of the prime necessities for life of each living organism on the earth. The pH level and turbidity dimension of water assumes natural job in surveying the nature of water. Water quality assumes natural job in the medical problems of individuals, plants and living beings on the earth. Especially, the primary wellsprings of water are rivers, waterfalls, and lakes. Downpour water running over the grounds contains numerous impurities and polluting influences that might be dissolvable or insoluble. The primary point is to gauge the pH level and turbidity level in the drinking water just as in the sewage water from ventures that are crashed into the waterfalls and furthermore the water utilized for horticulture.

The goals of the framework are given beneath

To sketch the analysis of the water quality monitoring system.

To evaluate the pH and turbidity parameters in the real time environment using the sensors.

To collect data from rural and less developed areas and store that data in the web page.

Thus the data collected is sent to the cloud with the help of wifi module present in the Arduino family microcontroller board. To exhibit the real time data on PC.

IV. Hardware Specification

The pH and turbidity parameters are used for checking the quality of water. These are calculated using the pH sensor and turbidity sensor which are connected to the Node MCU Microcontroller esp8266 as shown in Fig, 1. GSM module is a circuit that will be utilized to convey between a cell phone or registering framework and GSM.

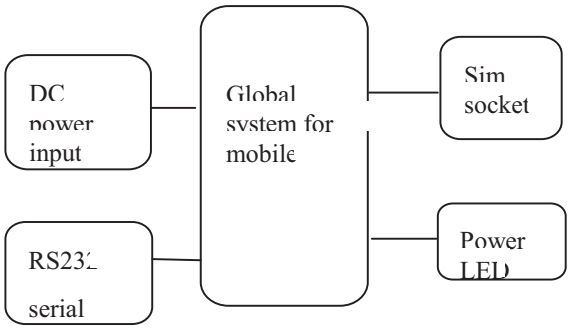


Fig 1. Block diagram for GSM

The SI unit for magnetic field strength H is A/m. However, if you wish to use units of T, either refer to magnetic flux density B or magnetic field strength symbolized as $\mu_0 H$. Use the center dot to separate compound units, e.g., "$A \cdot m^2$."

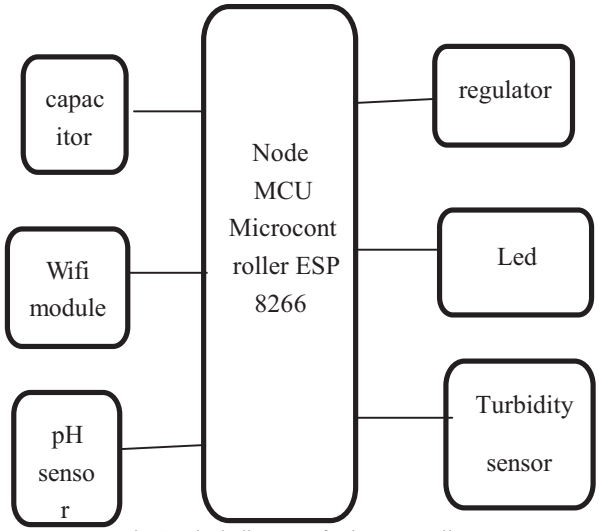


Fig. 2. Block diagram of Micro controller

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:16:13 UTC from IEEE Xplore. Restrictions apply.

A. pH Sensor

pH is the truncated type of potential hydrogen. The pH amplifier is used in combination with the pH sensor to calculate the acidity and alkalinity of the water or any other liquid. When kept in the water, the pH is calculated based on the movement of hydrogen and hydroxyl ions. The range of pH scale is from 0-14. The pH values range from 0-7 are considered as the acidic solutions. The pH of drinking water is 7, it is considered as neutral. The pH range from 7-14 is considered as alkaline.

B. Turbidity Sensor

The turbidity amplifier is used in combination with the turbidity sensor. The turbidity of water is calculated using turbidity sensor only. If the turbidity of the water is less than 60, then it is suitable for drinking, with some measures. The impurities which are invisible for human eye can also be detected using turbidity sensor in Fig. 3.

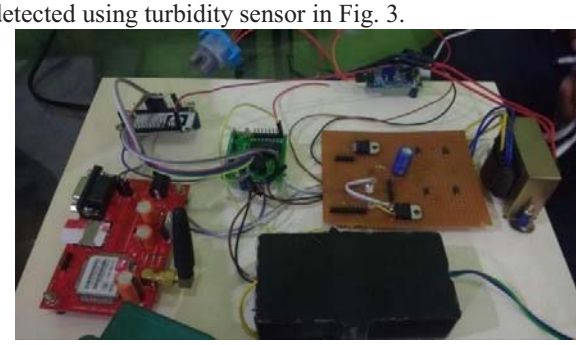


Fig. 3. Experimental setup

V. System Implementation

A. Existing System

The current Water Quality observing framework include human towards looking at the water Quality, Testing. At present many measures are taken and innovative development has been introduced in water quality checking. These are finished by utilizing automated fish, laser bar and advanced camera. Likewise enquiry about the fish has been done by utilizing remote sensors. Notwithstanding observing the water quality, restricted work is completed in applying AI strategy including the nature of water. The disadvantage of the current framework is that there is no completely robotized water Quality checking framework utilizing Sensors. Likewise framework does not have insight which takes into consideration dissecting the data for the forecast. These frameworks are worked for correspondence inside a little land zone.

B. Proposed System

In proposed system, the water quality will be checked by the Ph and turbidity sensors which are connected to the Node MCU microcontroller. The data is collected and is sent to the cloud for storage. The values are stored in a tabular format. We can get alert message notification to the registered mobile number through sms if the sensor reaches above the limit i.e Ph<7 and turbidity>60 using GSM. The values that are stored in a tabular form are analysed in the form of line graph and pie charts.

C. Advantages in Proposed System

The following are the advantages in the proposed system:

- Cloud storage.
- Efficient analysis in the form of bar graphs and pie charts.
- It can be constructed at minimum cost.
- Used in home applications.

VI. System Design

In this framework we present the structure of IOT Based Water Quality Monitoring System that screen the nature of water continuously which is shown in Fig. 4. It screens every 5 seconds and update the information. This framework comprises a few sensors which measure the water quality parameter, for example, pH, turbidity.

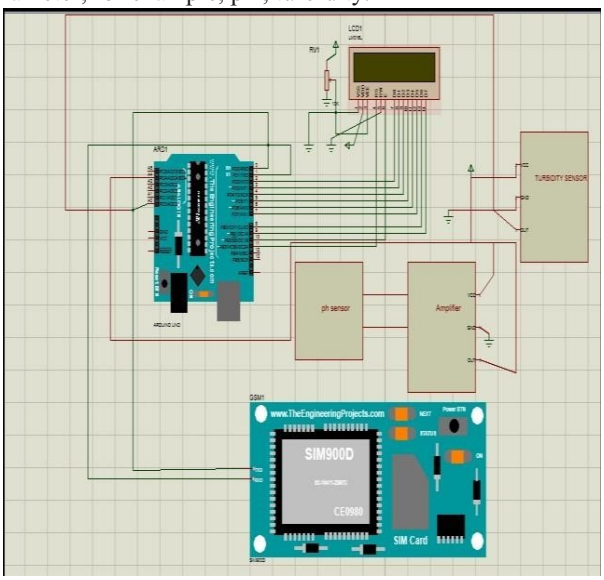


Fig. 4. Circuit Diagram

The deliberate qualities from the sensors are prepared by microcontroller to make the perfect for wifi module. These are transmitted remotely to the call controller, here it is the web page application, utilizing wifi convention. At last sensors information can be seen on the web browser application using the trending cloud computing.

VII. Analysis

The wifi module which is connected to the microcontroller update the sensors data continuously to the cloud. In the cloud it can be retrieved from anywhere. The web application with login credentials is created and the analysis of resultant values are declared in the form of line graph and pie chart. The pH values and turbidity values are shown as piechart in Fig. 5 and Fig. 6 respectively.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:16:13 UTC from IEEE Xplore. Restrictions apply.

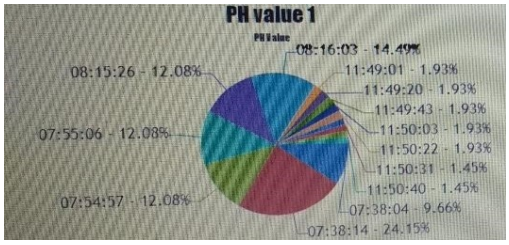


Fig. 5. pie chart for pH values

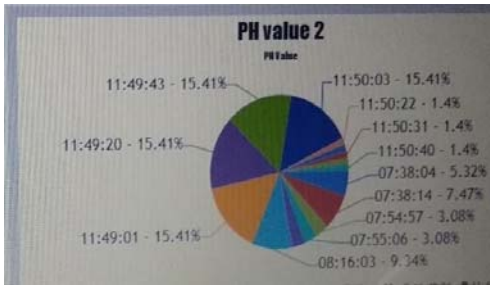


Fig. 6. Pie chart for turbidity values

VIII. Conclusion

The Project "An IoT based system for water quality monitoring" has been successfully designed and experimented. We have seen the success of sensors in various fields; the same idea has been applied to this water quality monitoring system. In this paper we have analysis different water quality monitoring systems to do the same. All these techniques are expensive and difficult in terms of analysis and collecting the data.

Acknowledgment

"This work has been carried out at DST-FIST sponsored Wireless Sensor Network and IOT Lab(sanction orderno. :SR/FST/ETI-413/2018 Dated:08th February, 2018), Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology."

References

[1] M.F. Saaid, A.Sanuddin, Megat Ali "Automated ph controller system for hydroponic cultivation" Symposium on computer applications and industrial electronics(ISCAIE) IEEE, pp.186-190,2015.

[2] Lakshmanan, L. and Jalasri., (2019), "A Survey: Integration of IoT and Fog Computing", IEEE International Conference on , ICCSP 2018, IEEE Conference Proceeding International Conference on Green Computing and Internet of Things (ICGCIoT),pp. 0235 - 0239, DOI: 10.1109/ICGCIoT.2018.8753010.

[3] Ayushi S.jaiswal, Vaidehi Baporikar "Embedded wireless data acquisition syatem for unnamed vehicle in underwater" IEEE underwater technology, vol 10,pp.1-6,2015.

[4] F Ntambi, C P kruger, B J silva "Design of a water management system" AFRICON IEEE pp.1-5,2015

[5] Jesudoss A., Muthuram B.O. and Lourdson Emmanuel A., "Safe Driving using IoT Sensors", conducted by International Conference on Engineering and Advancement in Technologies (ICEAT 2018), conduted by Sri Krishna College of Technology, Kovaipudur, Coimbatore

[6] Ji wang, Xiao-liRen, Yu-lishen "A Remote wireless sensor networks for water quality monitoring" International conference on innovative computing and communication, IEEE pp.7-12, 2010

[7] Cheng-Liang Lai, Chien-Lun chiu "Using image processing technology for water quality monitoring system" International conferenceonmachine learning and cybernetics.IEEE pp.1856-1861,2011.

[8] Lakshmanan, L. and Jesudoss and Ulagamuthalvi, (2018), "Cluster Based Routing Scheme for Heterogeneous Nodes in WSN–A Genetic Approach",ICICI 2018: International Conference on Intelligent Data CommunicationTechnologies and Internet of Things (ICICI) 2018,LNDECT, volume 26,pp 1013-1022

[9] Aravindan S.Rao, Stephan Martial, Jayavardhan Gubbi, "Design of low cost autonomous water quality monitoring system" International conference on advances in computing communications and informatics(ICACCI) IEEE,pp.14-19.2013.

[10] H Doukkali, L Nuaymi, "Analysis of MAC protocols of underwater sensor networking" Vehicular Technology conference UTC 2005 IEEE 61st,vol.2,pp.1307-1311,2005

[11] IF Akyildiz, D pompili, T Melodia,"Underwater acoustic sensor networks:research challenges", Ad hoc Networks,vol.3,no.3,pp.257-259,2005.

[12] Jesudoss A., R. Vybhavi and B. Anusha, "Design of Smart Helmet for Accident Avoidance", IEEE International Conference on Communciations and Signal Processing, organized by Adhiparasakthi Engineering College, Chennai, April 4 to 6, 2019.

[13] I.Vasilesu, K.Kotay, D.Rus, P.Corke "Data collection storage and retrieval with an underwater sensor network", IEEE Sensys, pp.154-165, 2005.

[14] Niel Andre Cleote, Reza Malekian and Lakshmi nair "Design of smart sensors for real time water quality monitoring" IEEE Access vol-4,no.9, pp.1-16,2014.

[15] P-Jiang, H.Xia, Z.He, Z-Wang "Design of a water environment monitoring system based on wireless sensor networks", Sensors,Vol.9,no.8,pp.6411-6434,2009.

[16] R.Fisher, L.Ledwaba, G.P Hancke, C.P Kruger, "Open hardware :A Role to play in wireless sensor Networks" sensrs 2015, Vol.15, no.3,pp 6818-6844,2015.

[17] T-P Lambrro, C.C Anastastiou, C.G Panayiotou, "A nephelometric turbidity systems for monitoring residential drinking water quality" in sensor Applications Experimentation Vol 29,2010.

[18] Vinod Raut, Sushama shelke, "Wireless acquisition system for water quality monitoring", Advaances in singleprocessing (CASP) conference on pp.371-374,2016.

[19] Yi Hong Wang, Ju Chen, Shi Feng Tao, "Waelet Fusion Application in water quality warning based on bio-detection technology",Applied mechanics ad materials, Vol 295-298,pp 924,2013.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 16:16:13 UTC from IEEE Xplore. Restrictions apply.

6

Analysis of skin lesions using machine learning techniques

J. Bethanney Janney[a], S. Emalda Roslin[b], and S. Krishna Kumar[a]

[a]Department of Biomedical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India, [b]Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

1 Introduction

Cancer develops due to abnormal cell growth having the possibility of spreading to other parts of the body. Skin cancer is one of the best-known cancer types and is widely seen, particularly in those with light-colored skin. The prevalent types of skin cancer are melanoma, basal cell carcinoma (BCC), and squamous cell carcinoma (SCC). Among these three, both basal and squamous, along with some other less common types, are classified as nonmelanoma skin cancer (NMSC). Basal cell carcinoma is less obtrusive, while squamous cell tumors tend to spread at a quicker rate than basal. More than 90% of skin cancers are triggered by the ultraviolet radiation of the sun. Exposure to radiation increases the risk for all three types of skin cancer. Other risk factors in addition to the sun are ionizing radiation (for example, X-rays, man-made radiation sources such as tanning beds), natural cancer-causing agents, and age. Tanning beds in particular are known to be one of the primary initiators of basal cell and squamous cell skin cancers. The World Health Organization currently places use of tanning beds in its highest hazard classification for skin disease. Between 20% and 30% of melanomas start from moles. Individuals with light skin are at greater risk of skin cancer [1, 2].

Melanoma is a dangerous form of skin cancer that occurs most commonly in people with less melanin pigment (Northern Europe and North America). It is caused by ultraviolet-A radiation (UV-A), which indirectly damages the DNA. The survival rate of persons affected is poor. Some lesions appear brown or black in color. Early detection and diagnosis of skin cancer can prevent it from spreading and becoming fatal [3]. The main objective of this research is to evaluate three classifiers—support vector machine (SVM), artificial neural network

Computational Intelligence and Its Applications in Healthcare. https://doi.org/10.1016/B978-0-12-820604-1.00006-6
© 2020 Elsevier Inc. All rights reserved.

(ANN), and Naïve Bayes—to have a better understanding of their performance in terms of detection of skin cancers. The work described in this chapter involves manual cropping of preprocessed dermoscopic images and extracting features from them; this is then given as input to the SVM, ANN, and Naïve Bayes classifiers. The performance of these classifiers is then analyzed based on accuracy, sensitivity, and specificity, to determine their relative effectiveness [4, 5].

2 Related works

Sumithra et al. developed a computer-aided diagnosis system (CADS) using shade and surface attributes depicting the derived lesion areas; SVM and k-nearest neighbor (k-NN) classifiers were used to classify a dataset composed of five distinct classes of diseases, and the authors analyzed the outcomes of the classification [6]. Tan et al. presented a genetic algorithm (GA) to determine the most selective subset of characteristics from a skin region's structure, shade, and surface characteristics. An SVM classifier was utilized to recognize benign and cancerous tumors [7].

Premalatha and Ravichandran aided dermatologists in diagnosing cancer in its early phases by creating a CADS to classify and predict melanoma skin lesions. Fifteen shape and statistical characteristics were derived from the segmented images and used through the assistance of separate machine-learning algorithms [8]. Odeh and Baareh have attempted to create a coherent diagnostic system of classical k-NN with sequential scanning feature selection, classical k-NN with complex GA, complex ANN with complex GA, and an optimized method of neuro-fuzzy inference to determine the exact choice-making process in human biomedicine [9]. Sui et al. proposed the technique of texture feature extraction for classification of melanoma. This diagnostic approach makes use of image analysis and artificial intelligence [10]. Patil et al. suggested a study on controlling malignant melanoma with a noninvasive ABCD technique utilizing image analysis through MATLAB software. The image procurement was for identification and elimination, segmentation of lesions, extraction and classification of attributes, which is included in an automatic image analysis module [11].

DeVries and Ramachandram proposed a method using advanced multiscale convolutionary neural models for the analysis of skin lesions. This technique used an immediate posttrained Inception-v3 network on the ImageNet database tuned to identify skin lesions using two separate sample object levels [12]. Zhang et al. used an optimal convolutionary neural network (CNN) to detect skin cancer early. To optimize the CNN, the improved whale optimization algorithm was used. To analyze the proposed method, it was contrasted using two different datasets with certain different approaches [13].

Nasr-Esfahani et al. have been engaged in developing an extremely complicated technique of deep learning to help experienced dermatologists. Images with brightness and clamors are further preprocessed to reduce these artifacts. Educated by a huge number of learning sets, the CNN algorithm distinguishes melanoma from benign instances [14].

3 Materials and methods

The images used in this study were dermoscopic images. Initially the images were segmented by assigning values to the region to be cropped based on the region of concern using manual cropping. Extraction of the gray level co-occurrence matrix (GLCM), texture, and wavelet features was also carried out. Carcinogenic (malignant) and noncarcinogenic (benign) skin lesion input images were collected and analyzed as a database; 430 dermoscopic images, including 320 benign and 110 malignant lesions, were evaluated by the proposed system. Three classifiers were used in the work, i.e., SVM, ANN, and Naïve Bayes. Block diagrams showing classification of skin cancers using SVM, Naïve Bayes, and ANN classifiers are presented in Fig. 1.

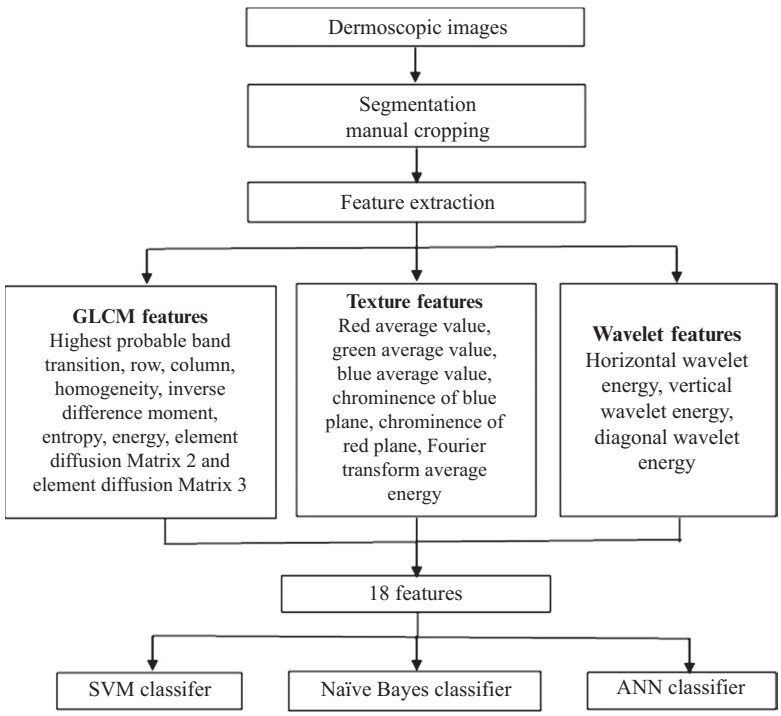


Fig. 1 Block diagram of classification of skin lesions using SVM, ANN, and Naïve Bayes classifier.

3.1 Research database

The input images were obtained from the certified database of International Skin Imaging Collaboration (ISIC Archive), in which the images of benign and malignant skin lesions are collected and stored [15].

3.2 Skin lesion segmentation

Segmentation of the image was accomplished by separating a digital picture into different segments. The objective of this separation is to improve and alter the description or segregate the essential portion of an image so it can be analyzed rapidly. Here, the images were assessed and segmentation was done by manual cropping, which enables the individual to choose the affected area (i.e., lesion) from images. Fig. 2 shows the segmentation output of 10 sample images.

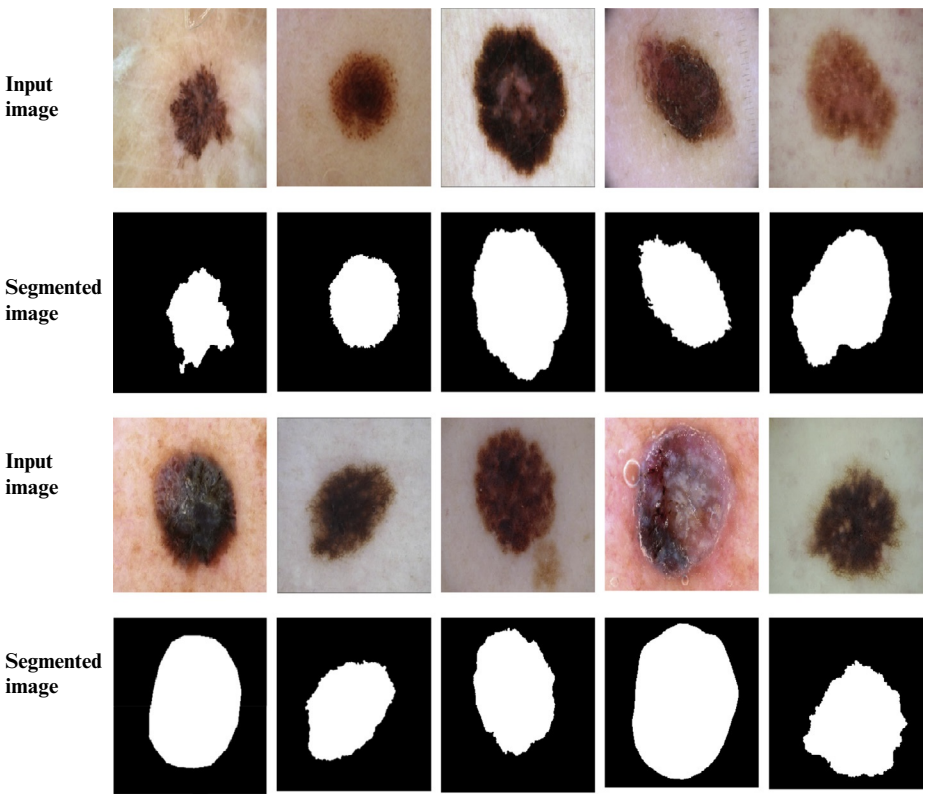


Fig. 2 Segmentation output.

3.3 Feature extraction

An original collection of observed data is used to extract features and it forms derivative variables (features) that hold the information needed to make training effective and generalized. In other words, the understanding or interpretation of data is much easier and better. It reduces the dimensions of the given images. The analysis of an image having a large number of variables makes it more difficult, as it requires more computation power and memory.

This might lead to an overfit in classification and training steps and hence result in poor sample classification. Feature extraction enables the construction of different patterns of the variables that describe the data with adequate accuracy, without causing any other changes in the information gathered from the image [16].

GLCM attributes, texture attributes, and wavelet attributes were derived from the segmented images [17]. On the whole, 18 features were extracted. These features are explained in the following sections.

3.3.1 GLCM features

The GLCM looks at two pixels and the relation between them, which is the second-order texture. The co-occurrence matrix contains the gray value relationship given by a kernel mask. The neighboring pixels in the transformation are defined by certain directions (e.g., 0, 45, 90 degrees, etc.). The direction can also be in a negative direction (reverse). The position depends on the similarity in gray level values [18, 19].

(a) Highest probable band transition: The maximum value is selected from the GLCM matrix. The max () command selects the highest value from the matrix and displays it.

(b) Row: The highest values in the row are depicted in this command and this is done by first obtaining the maximum values from the whole image.

(c) Column: In the column command, the highest values in the column are depicted and this is done by first obtaining the maximum values from the whole image.

(d) Homogeneity: This refers to uniformity of the pixels in the given image, and it shows the comparative arrangement of equal pixel values.

(e) Inverse difference moment: IDM uses a variety of values to define whether the image is textured or not. If the local gray level is uniform and the reverse GLCM is large, the IDM value will be large.

$$f = \sum_{i=0}^{N-1} \sum_{j=0}^{n-1} \frac{1}{1+(i-j)^2} P_{d,\beta}(i,j) \tag{1}$$

(f) Entropy: The measure for determining uncertainty, entropy is also used to depict the texture of the image. If all the co-occurrence vector components are the same, the entropy's magnitude will be a maximum.
(g) Total energy: This value is used to measure the uniformity of the image. When all the pixels are the same, the energy rate is large. The energy value of a constant image is 1.
(h) Element diffusion Matrix 2, and element diffusion Matrix 3: Element difference moment of order k: $P_i\ P_j\ (i-j)\ k\ c_{ij}$; in general this indicator has small qualities when large C estimates are close to diagonal rule. For this role administrator, large qualities close to primary corner to corner would show that groups of consistent power running "1 pixel to one side and 1 down" are likely.

3.3.2 Texture features

The appearance of an object and its surface characteristics together make up texture. The essential step of collecting such characteristics for the texture analysis procedure is called extraction of the texture attribute. It is a core step because of the significant information obtained from the images [20].

The texture features are:

(a) RGB color scale: The RGB color system is an integrated hue model that reproduces a wide range of colors by integrating red, green, and blue light together.
(b) Red mean value: The color model's correlation coefficients of pixels in the red plane are obtained.
(c) Green mean value: The color model's correlation coefficients of pixels in the green plane are obtained.
(d) Blue mean value: The color model's correlation coefficients of pixels in the blue plane are obtained.
(e) YCbCr color space: Efficient representation of images can be obtained using the YCbCr color space, by splitting the image's illumination and chromaticity. In other words, YCbCr is used in information systems to technically encode color data.
(f) Chrominance of blue plane (Cb): This gives the color information in the blue plane of the YCbCr color space.
(g) Chrominance of red plane (Cr): This gives the color information in the red plane of the YCbCr color space.
(h) Fourier transform average value: The frequency domain is an important part of machine vision, for the decomposition of an input into its sine and cosine components. The discrete Fourier transform (DFT) is the sampled Fourier transform which converts a finite sequence of equally spaced functions into equally

spaced discrete time Fourier transform. The fast Fourier transformation (FFT) is a DFT algorithm that reduces the number of computations.

$$F(k,l)=\sum_{i,j=0}^{N-1} f(i,j)\mathrm{e}^{-i2\pi\left(\frac{ki}{N}+\frac{lj}{N}\right)} \tag{2}$$

3.3.3 Wavelet feature extraction

Multiresolution images are generally represented and analyzed using wavelets. Wavelet transformation is equivalent to the Fourier transform and is widely used in noise removal in image compression.

(a) Horizontal wavelet energy: The main detailed picture of parallel elements, consisting of variations in brightness along the horizontal plane of the actual image determined by the 2D wavelet transform.

(b) Vertical wavelet energy: The main detailed picture of perpendicular elements, consisting of variations in brightness along the vertical plane of the actual image determined by the 2D wavelet transform.

(c) Diagonal wavelet energy: The main detailed picture of adjacent elements, consisting of variations in brightness along the diagonal plane of the actual image determined by the 2D wavelet transform.

3.4 Classification

Classification is the final phase of the detection system. It includes the method used for classifying the given sample into carcinogenic (malignant) and noncarcinogenic (benign) images. When separating the necessary features from the given image, the correct results are used for the purpose of identification. Many different classifiers can be used to perform the classification process. MATLAB software was used here for classification purposes and a computer-based ANN classifier, SVM classifier, and Naïve Bayes classifier were implemented in the software. Depending on the classification outcomes, the accuracy, sensitivity, and specificity of all three classifiers were obtained. Finally, based on the quality of the three classifiers, a quantitative analysis was carried out [21].

Support vector machine (SVM) is a mathematical computational model, like a neural network, which builds hyperplanes that specify classification decision boundaries. The SVM determines the criteria for finding a decision boundary by maximizing the distance between the categorized labels. SVM can be applied to binary classification and classification of multiple classes. The original SVM is

primarily used for linear classification, but the kernel trick can be used to perform nonlinear classification. SVM is based on statistical learning and utilizes a supervised learning model. An unlimited number of probable decision boundaries exist, but SVM chooses the best decision boundaries; the related hyperplane is formulated precisely as

$$\text{H1}: \quad x_i * w + b \geq 1 \quad \text{if } y_i = +1(d+) \tag{3}$$

$$\text{H2}: \quad x_i * w + b \leq 1 \quad \text{if } y_i = -1(d-) \tag{4}$$

where x_i = vector feature, y_i = category label, w = weight, b = bias rate, $d+$ = closest positive point range, $d-$ = closest negative point range. H1 and H2 are support for hyperplanes. The plane points H1 and H2 are vectors of support. Limits must be greatly increased in order to obtain the best outcomes. Therefore, the problem of optimization can be described using Eq. (5). At each iteration, values are reorganized until the convergence approaches [22]:

$$\psi = \min\left(\frac{1}{2} w^T w\right) \tag{5}$$

The dermatoscope images were utilized with the ANN classifier to obtain a classifier result. Subsequently, there were 18 features, with the classifier containing 18 inputs. The number of hidden nodes used here was 10 with 1 output neuron. A logistic sigmoid activation function was used. The network capacity can be either 0 or 1, with 0 implying a healthy state and 1 implying a cancerous state. This classifier was designed using MATLAB code. To train the ANN, a back propagation network (BPN) was used [23, 24].

The Naïve Bayes classifier makes use of the idea introduced by Thomas Bayes (1702–61) in the Bayes theorem. It is focused on supervised training and utilizes probability to assess the category to which an object belongs. After normalization of the extracted values, the values are separated into two classes. Using a Gaussian probability density function (GPDF), the probability of the image can then be obtained [25].

4 Results and discussion

The dataset of images obtained from an authorized skin cancer image database was segmented using manual cropping. The features were extracted and stored in an Excel spreadsheet. These features are GLCM, texture, and wavelet features, which were classified using SVM, ANN, and Naïve Bayes classifiers. The suggested system was tested on 430 dermoscopic images containing 320 benign and 110 cancerous lesions.

The efficiency outputs of the three classifiers were measured and the most effective among the three were calculated based on precision, specificity, and sensitivity, and then further evaluated utilizing an ROC plot. The 18 extracted elements of the benign and malignant images are reported in Tables 1 and 2.

4.1 Performance analysis

The classifier's reliability was determined with regard to precision, specificity, and sensitivity using a confusion matrix [26]. These parameters can be evaluated using the following formula:

$$\text{Accuracy} = \left[(\text{TP} + \text{TN}) \times 100 / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \right] \tag{6}$$

$$\text{Sensitivity} = \left[\text{TP} \times 100 / (\text{TP} + \text{FN}) \right] \tag{7}$$

$$\text{Specificity} = \left[\text{TN} \times 100 / (\text{TN} + \text{FP}) \right] \tag{8}$$

where:

TP-true positive: the malignant image being assessed accurately as malignant

TN-true negative: the benign image being assessed accurately as benign

FP-false positive: the benign image being wrongly assessed as malignant

FN-false negative: the malignant image being wrongly assessed as benign.

The output of the performance analysis in terms of percentages is shown in Table 3.

It can be inferred from Table 3 and Fig. 3 that the ANN classifier has the highest accuracy, sensitivity, and specificity when compared with SVM and the Naïve Bayes classifier. Hence it was found to be the most suitable classifier and further analysis was carried out based on the ROC curve [27, 28].

Since two parameters are considered, the receiver operating characteristic curve (ROC) can be used for analyzing binary classifiers. The ROC is calculated by using parameters to map the true positive level against the false positive level. The true positive level is often known as sensitivity, and the false positive level is known as specificity. Fig. 4A–C displays the ROC plots of the three classifiers. It can be concluded that the ROC plot shows the trade-off between sensitivity and specificity (any increase in sensitivity is preceded by a reduction in specificity). The more the curve crosses the left boundary, and then the ROC space's top edge, the more accurate the measurement is. The closer the curve comes to the ROC space's 45-degree diagonal, the less exact the test. Thus the ANN classifier tends to be the highest, followed by the Naïve Bayes and then the SVM classifiers.

Table 1 Feature extraction results of benign images.

Image no.	r-Plane	b-Plane	g-Plane	Cb plane	Cr plane	Average of frequency domain feature	Highest probable band transition	Row
B1.	0.1896	0.1624	0.1836	0.5072	0.5827	7.3486	0.8837	2
B2.	0.3563	0.2267	0.1724	0.4591	0.5628	6.9754	0.5204	2
B3.	0.4914	0.2814	0.2026	0.4366	0.5998	6.9599	0.4256	3
B4.	0.4430	0.2287	0.1435	0.4330	0.6021	7.3512	0.5887	3
B5.	0.4151	0.2303	0.1663	0.4467	0.5878	6.7538	0.4250	2
B6.	0.4235	0.2451	0.1759	0.4421	0.5637	7.6763	0.5145	3
B7.	0.4728	0.2264	0.1658	0.4291	0.5862	6.9218	0.4869	2
B8.	0.4940	0.2191	0.1721	0.4324	0.5828	7.4336	0.4321	3
B9.	0.4285	0.2231	0.1521	0.4421	0.5962	6.3129	0.4783	3
B10.	0.4657	0.2372	0.1856	0.4357	0.5654	6.9434	0.4198	2

r-plane, red plane; *b-plane*, blue plane; *g-plane*, green plane; *Cb plane*, chrominance of blue plane; *Cr plane*, chrominance of red plane; *IDM*, inverse difference moment, *EDM2*, element diffusion matrix 2; *EDM3*, element diffusion matrix 3; *HEN*, horizontal wavelet energy; *VEN*, vertical wavelet energy; *DEN*, diagonal wavelet energy.

Table 2 Feature extraction results of malignant images.

Image no.	r-Plane	b-Plane	g-Plane	Cb plane	Cr plane	Average of frequency domain feature	Highest probable band transition	Row
M1.	0.5244	0.3547	0.3826	0.5072	0.6053	40.3486	0.5837	4
M2.	0.4892	0.3167	0.3724	0.5091	0.6328	30.9754	0.5204	4
M3.	0.4974	0.3914	0.3021	0.5266	0.6954	36.9599	0.5216	3
M4.	0.4724	0.4215	0.3945	0.5110	0.6021	41.3512	0.5887	3
M5.	0.4671	0.3903	0.3153	0.5237	0.6721	44.7538	0.4250	6
M6.	0.5526	0.3565	0.1759	0.4231	0.6378	34.6763	0.7159	3
M7.	0.4728	0.3608	0.3650	0.4873	0.5510	163.9218	0.5391	4
M8.	0.4940	0.3952	0.2028	0.4324	0.5958	737.4336	0.5489	3
M9.	0.4285	0.4701	0.3331	0.5060	0.5670	26.3129	0.3468	3
M10.	0.6299	0.4667	0.4895	0.4878	0.5720	12.9434	0.3635	6

r-plane, red plane; *b-plane*, blue plane; *g-plane*, green plane; *Cb plane*, chrominance of blue plane; *Cr plane*, chrominance of red plane; *IDM*, inverse difference moment, *EDM2*, element diffusion matrix 2; *EDM3*, element diffusion matrix 3; *HEN*, horizontal wavelet energy; *VEN*, vertical wavelet energy; *DEN*, diagonal wavelet energy.

Column	Homogeneity	IDM	Entropy	Energy	EDM2	EDM3	HEN	VEN	DEN
2	0.9845	0.9763	0.7791	0.4843	0.0473	−0.00079	0.7633	0.6553	0.0843
2	0.9642	0.9699	1.6792	0.3984	0.0600	−0.00144	1.3656	0.5355	0.1097
3	0.9660	0.9723	1.8832	0.3273	0.0553	−0.00132	0.4846	0.3571	0.0553
3	0.9516	0.9590	1.6913	0.4203	0.0818	−0.00122	0.9904	0.7058	0.0718
2	0.9660	0.9712	1.8223	0.3555	0.0575	0.00195	0.4386	0.4789	0.0412
3	0.9845	0.9618	1.6546	0.4728	0.0564	−0.00137	0.8638	0.3223	0.0535
2	0.9902	0.9653	1.6351	0.3588	0.0568	−0.00187	1.2698	0.5162	0.0684
3	0.9754	0.9624	1.6987	0.4186	0.0581	−0.00129	0.7573	0.6271	0.0429
3	0.9762	0.9607	1.7854	0.3935	0.0502	−0.00138	0.9075	0.4997	0.0627
2	0.9812	0.9524	1.7133	0.3676	0.0587	−0.00179	1.2530	0.7156	0.0475

Column	Homogeneity	IDM	Entropy	Energy	EDM2	EDM3	HEN	VEN	DEN
4	0.9845	0.9757	1.9891	0.7843	0.0473	−0.00278	0.6433	1.2410	0.0673
4	0.9842	0.984	1.8592	0.6584	0.0400	−0.00244	1.7356	1.2101	0.0697
3	0.9875	0.9723	1.8832	0.7273	0.0473	−0.00243	0.9246	1.3571	0.0655
3	0.9831	0.9754	1.9913	0.5903	0.0432	−0.00222	0.9904	0.6058	0.0612
6	0.9810	0.9725	1.9223	0.6455	0.0575	0.00395	0.8786	1.4789	0.0634
3	0.9652	0.9718	1.9926	0.5472	0.0564	−0.00033	0.8638	1.2235	0.0635
4	0.9609	0.9689	1.9731	0.3583	0.0620	−0.00205	0.6698	0.4160	0.0684
3	0.9706	0.9759	1.6436	0.4127	0.0481	−0.00019	0.4773	0.6271	0.0629
3	0.9762	0.9798	2.2979	0.2635	0.0403	0.000585	0.5075	0.4997	0.0627
6	0.9667	0.9734	2.7133	0.2076	0.0532	−0.00179	1.9508	3.3316	0.0745

Table 3 Performance analysis for different classifiers.

Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
ANN	**89**	**90**	**88**
Naïve Bayes	71	90	76
SVM	87	78	72

ANN, artificial neural network; *SVM*, support vector machine.

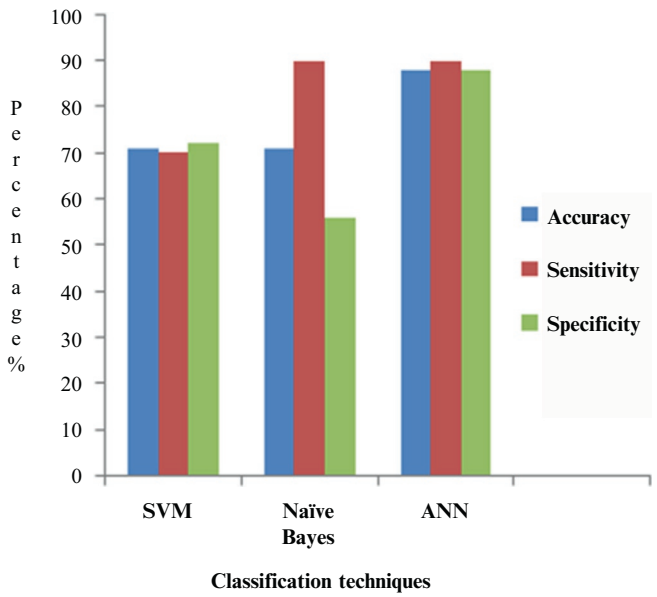


Fig. 3 Performance analysis of classifiers.

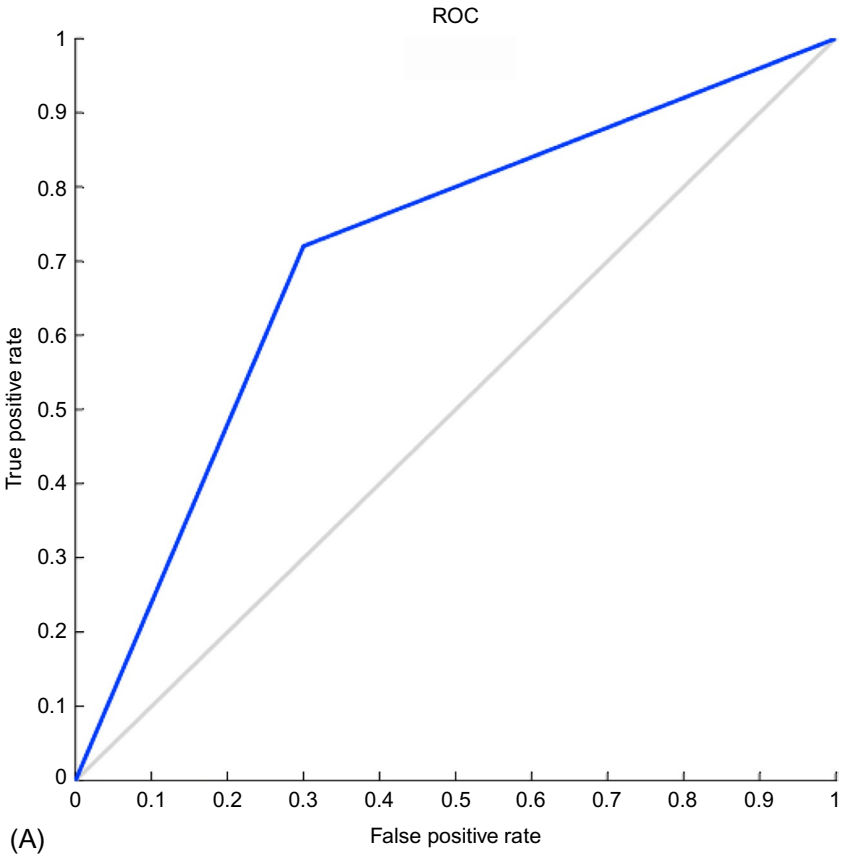


Fig. 4 (A) ROC plot of SVM classifier.

(Continued)

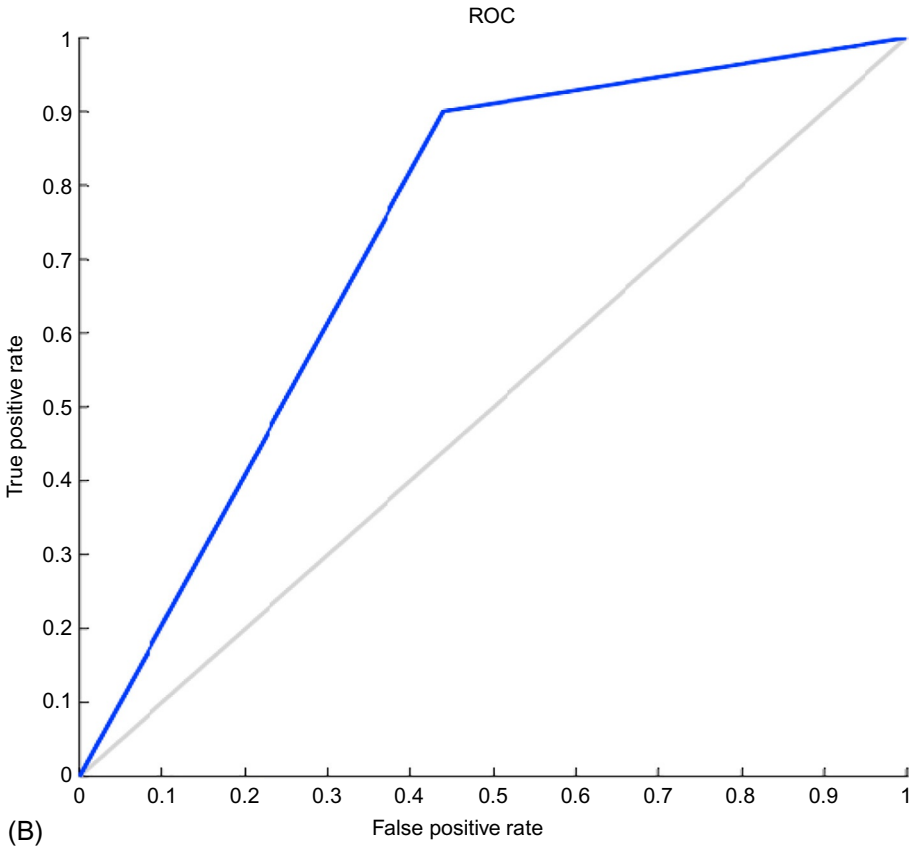


Fig. 4, cont'd (B) ROC plot of Naïve Bayes classifier.

(Continued)

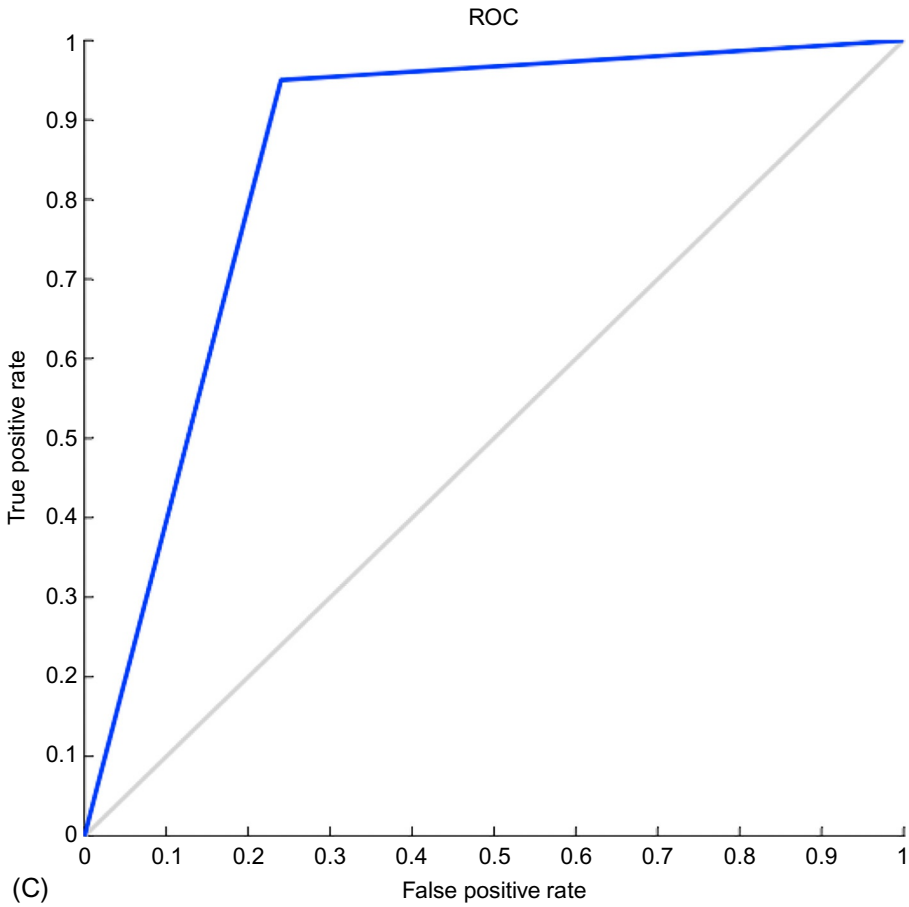


Fig. 4, cont'd (C) ROC plot of ANN classifier.

5 Conclusion

Three algorithms were analyzed in this work, depending on supervised learning. The images were categorized into two classes: noncarcinogenic (benign) and carcinogenic (malignant). The attributes from the images were extracted utilizing a GLCM framework and texture and wavelet attributes. The derived attributes were fed into the architectures of the classifiers and the output evaluation dependent on precision, sensitivity, and specificity were determined. The SVM's precision (71%) sensitivity (70%), and specificity (72%) were less compared to Naïve Bayes, with an accuracy of 71%, sensitivity of 90%, and specificity-56%. Hence, it is concluded that the ANN classifier had the best performance among the three, with an accuracy of 89%, sensitivity of 90%, and specificity of 88%. Therefore, to identify benign and malignant skin cancer, ANN can be used as an effective classifier. Therefore, the described computer-based skin lesion detection system illustrates an automatically generated screening technique for accurate and reliable evaluation of skin carcinoma with extraordinary potential for noninvasive treatment approaches.

Acknowledgment

We, the contributors, wish to thank the management of Sathyabama Institute of Science and Technology for exemplary guidance, encouragement, and resources offered to us.

Conflict of interests

The authors confirm there is no conflict of interest in publishing this chapter.

References

[1] R. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2016, CA Cancer J. Clin. 66 (1) (2016) 7–30.
[2] American Cancer Society, Melanoma Skin Cancer, 20 March 2015.
[3] N. Cheerla, D. Frazier, Automatic melanoma detection using multistage neural networks, Int. J. Innov. Res. Sci. Eng. Technol. 3 (2) (2014) 9164–9183.
[4] A. Santosh, G. Sadashivappa, Statistical analysis of skin cancer image—a case study, Int. J. Electr. Commun. Eng. 3 (3) (2014) 1–10.
[5] J. Kamalakkannan, S. Bhavani Sankari, Detection of malignant skin cancer based on automated image analysis and classification, Int. J. Innov. Trends Emerg. Technol. 1 (1) (2015) 26–37.
[6] R. Sumithra, M. Suhil, D.S. Guru, Segmentation and classification of skin lesions for disease diagnosis, Procedia Comput. Sci. 45 (2015) 76–85.

[7] T.Y. Tan, L. Zhang, M. Jiang, An intelligent decision support system for skin cancer detection from dermoscopic images, in natural computation, in: Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 12th International Conference on IEEE, 2016, pp. 2194–2199.
[8] J. Premalatha, K. Ravichandran, Novel approaches for diagnosing melanoma skin lesions through supervised and deep learning algorithms, J. Med. Syst. 40 (96) (2016) 1–12.
[9] S.M. Odeh, A.M. Baareh, A comparison of classification methods as diagnostic system: a case study on skin lesions, Comput. Methods Prog. Biomed. 137 (2016) 311–319.
[10] H. Sui, M. Samala, D. Gupta, N. Kudu, Texture feature extraction for classification of melanoma, Int. Res. J. Eng. Technol. 5 (3) (2018) 1026–1029.
[11] M.R. Patil, A. Ghonge, M. Dixit, D. Joshi, Non-invasive ABCD monitoring of malignant melanoma using image processing in MATLAB, Int. Res. J. Eng. Technol. 4 (3) (2017) 1074–1081.
[12] T. DeVries, D. Ramachandram, Skin lesion Classification using deep multi-scale convolutional neural networks, Comput. Vis. Pattern Recognit. (2017). arXiv: 1703.01402v1.
[13] N. Zhang, Y.-X. Cai, Y.-Y. Wang, Y.-T. Tian, X.-L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network, Artif. Intell. Med. 102 (101756) (2020) 1–7.
[14] E. Nasr-Esfahani, S. Samavi, N. Karimi, et al., Melanoma detection by analysis of clinical images using convolutional neural network, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2016, pp. 1373–1376.
[15] ISIC, ISIC Archieve: The International Skin Imaging Collaboration: Melanoma Project, ISIC, 5 Jan 2016. Available from: https://isic-archive.com/#, (Accessed January 20, 2018).
[16] M. Ruela, C. Barata, J.S. Marques, J. Rozeira, A system for the detection of melanomas in dermoscopy images using shape and symmetry features, Comput. Methods Biomech. Biomed. Eng. Imaging Vis. 5 (2) (2017) 127–137.
[17] J. Bethanney Janney, S.E. Roslin, Classification and detection of skin cancer using hybrid texture features, Biomedicine 37 (2) (2017) 214–220.
[18] R. Goel, S. Singh, Skin cancer detection using GLCM matrix analysis and back propagation neural network classifier, Int. J. Comput. Appl. 112 (9) (2015) 40–47.
[19] L.-s. Wei, Q. Gan, T. Ji, Skin disease recognition method based on image color and texture features. Comput. Math. Methods Med. (2018) 1–10, https://doi.org/10.1155/2018/8145713.
[20] M. Mai, S. Mariam, A. Sharawy, Automatic detection of melanoma skin cancer using texture features, Int. J. Comput. Appl. 42 (20) (2012) 22–26.
[21] K. Nabin, M. Mishraa, E. Celebib, An overview of melanoma detection in dermoscopy images using image processing and machine learning, Comput. Vis. Pattern Recognit. (2016) 1–15. ArXiv: 1601.07843.
[22] Immagulate, S.M. Vijaya, Categorization of non-melanoma skin lesion diseases using support vector machine and its variants, Int. J. Med. Imaging. 3 (2) (2015) 34–40.
[23] T. Kanimozhi, A. Murthi, Computer aided melanoma skin cancer detection using artificial neural network classifier, Singaporean J. Sci. Res. 8 (2) (2016) 35–42.
[24] J. Albin, N. Muthu, E. Roslin, Text independent speaker recognition system using back propagation network with wavelet features, in: IEEE International Conference on Communication and Signal Processing, ISBN: 978-1-4799-3357-0, 2014, pp. 942–946.

[25] D.E. Chavan, K.V. Kale, Melanoma detection in dermoscopic images using Naïve Bayes classification algorithm, Int. J. Res. Technol. Stud. 5 (1) (2017) 1–5.
[26] E. Almansour, M. Arfan Jaffar, Classification of dermoscopic skin cancer images using color and hybrid texture, Int. J. Comput. Sci. Netw. Secur. 16 (4) (2016) 135–139.
[27] K. Hajian-Tilaki, Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation, Caspian J. Intern. Med. 4 (2) (2013) 627–635.
[28] D. Rofman, G. Hart, M. Girardi, J.K. Christine, J. Deng, Predicting non-melanoma skin cancer via a multi-parameterized artificial neural network, Sci. Rep. 8 (2018) 1–7.

Analysis Of Brain Tumor Segmentation using Convolutional Neural Network with Magnetic Resonance Imaging

B.Keerthi Samhitha, P.Yaswanth, Yellappa Gari Srinivasulu, Suja Cherukullapurath Mana and Jithina Jose

***Abstract*—Computer vision plays a vital role in human health care area. The various applications of these techniques have been playing a key role in reducing physicians judgment in diagnosis. Especially, in the case of brain related cases, it is suggested and required to take utmost care and a small error in diagnosis leads to disastrous. Although various methods like X ray, ct scan are available, mri is more reliable and danger free. High contrast images of brain tissues can be taken by using mri technique. Brain Tumor segmentation involves separation of tumors from normal brain tissues. The segmentation process helps in fetching required information for the purpose of diagnosis and what treatment has to be taken. It plays a vital role in medical image processing. The process involves risk because of abnormal form and irregularities at the boundaries. Determination of these tumors at the earlier stage helps to give better treatment to the patients. Usually, Physicians deal with the affected people by performing manual way of segmentation which results in the absence of time management. To overcome this manual segmentation problem, we developed CNN method of segmentation procedure which performs well by taking various factors like accuracy, recall and precision into consideration. MRI is one of the conservative techniques for prior diagnostic tool of brain tumor. The necessary measures for diagnosis and treatment planning have to be taken place quickly if the tumor is malignant in order to enhance successful treatment. By applying the following method, the doctors are capable of making accurate decision and required treatment can be done in an effective manner. The computer aided diagnostic system, MRI helps in reduction of workload by doctors.**

***Index Terms*—Convolution, Preprocessing, Matlab, Classification, Biomedical imaging, Kernel, Tumors, MRI, Segmentation.**

B.Keerthi Samhitha, is working as Assistant Professor in Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology, Chennai-119, India (phone: 9498331643, e-mail: samhitha711@gmail.com)

P.Yaswanth and Yellappa Gari Srinivasulu are UG Student in the Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology, Chennai-119, India (e-mail: yaswanthp.999@gmail.com, srinivassrini2018@gmail.com)

Suja Cherukullapurath Mana and Jithina Jose are working as Assistant Professor in Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology, Chennai-119, India (e-mail: cmsuja@gmail.com, jithinajose@gmail.com)

I. Introduction

A brain tumor refers to the group of abnormal cells that occur within the brain. They are of two types-benign and malignant. Benign tumors are not that much dangerous and can be distinguished easily from normal brain tissues. Malignant tumors are dangerous and they cannot be differentiated from normal brain tissues. In recent years, more cases are noticed regarding malignant tumors. Magnetic Resonance Imaging is a noninvasive medical technique mostly used because of high resolution, accuracy and contrast in obtaining an image. Image segmentation plays an important role in this process where a vast number of methodologies with different algorithmic techniques are implemented.

Image segmentation is the process of partitioning a digital image into disjoint, meaningful regions. The procedure involving segmentation helps for recognizing object, estimating boundary, compressing etc. Generally, two methods are involved viz., Hand-crafted featured and classifier method, which uses algorithms like random forest and support vector machine where the input is features given to classifiers and the other method is using CNN where the features are involved with training data. In this paper, we implemented Convolutional neural network methodology for brain tumor segmentation in MRI images. Our proposed method extracts features from pixel images with a minimum amount of preprocessing. Training time is reduced in proportion to number of parameters [1-22]. It involves composition of data layers, covered layers and yield layer. Feature Extraction is automatically done for the given task. High Efficiency is the most featured advantage of CNN.

The remainder of the paper is organized as follows: Section II defines the literature review. Proposed methodology is defined in section III. Section IV-VII describes the system archietecture, Image preprocessing, image segmentation, image classificationThe paper is concluded in section VIII.

II. Literature Review

[1].Konstantinos (2017) proposed brain segmentation by using 3d convolutional neural network. His method deals with taking an input image in multiple scales. It is an efficient method since the required numbers of parameters are low and operational cost is also low. The main aim was to separate lesions affected with injuries, stroke and tumor.

978-1-7281-4988-2/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:38:32 UTC from IEEE Xplore. Restrictions apply.

[2].Hussian (2017) developed a segmentation algorithm which takes the deep neural network algorithm into consideration to identify the tumor which doesn't have correct shape .The solution for overfitting is also found by using layers in patch processing. The Survival percentage of patients is increased and this algorithm also removes the unwanted noise as a part of the preprocessing method.

[3].Dong (2017) proposed non invading method of segmentation without ionization of radiation. The author proposed u-net based deep neural network which takes BRATS datasets for their implementation.

[4].Padole & Chaudhari [2016] proposed component analysis of detecting the tumors of brain. It uses mean shift and ncut algorithm for implementation. The feature extraction is difficult because of inter dependencies in the inner tissues of the brain.

[5].Abdel-Maksoud [2016] combined k means clustering algorithm and fuzzy c means method for the purpose of division. The various threshhold stages of segmentation at various levels of the process helps in accurate tumor detection. This method involves advantages from two algorithms. Accuracy is increased because of using fuzzy c means algorithm and the overall calculating time is reduced because of using k-means clustering algorithm.

[6].Menze [2010] has taken a fixed set of glioma patients and applied various algorithmic procedures for detection process. It was found good when all the regions are differentiated and it was not suitable for each sub region.

[7].A Selvapandian, [2016] believed that it was highly difficult for him to perform the detection of tumor regions because of low sensitivity in boundary regions. He applied neuro fuzzy approach for classification and training of extracted features. He used dataset of BRATS for assessing the performance.

[8].Solmaz Abbasi, [2010] developed tumor detection of brains by making use of 3 dimensional images. For the purpose of processing, coordinating Histogram and amending bias field have been used. For extracting the region of interest, otsu algorithm was applied. The region segmentation was done using random forest algorithm. Because of this method, feature extraction was done in a less span of time.

[9].SANJIV Kumar [2012] presented a cross breed methodology for the purpose of detecting brain tumor. Discrete wavelength transform technique was utilized to get the properties in classifying technique and here Root Mean Square is reduced and accuracy is increased.

III. Proposed Methodology

A. Matlab

Matlab is an effective language in case of computational purposes. It includes visualization, computing and development in a clear surroundings in which queries and consequences are expressed in familiar notation. It makes use in Data analysis, exploration, visualization, modeling, simulation and prototyping.

Convolutional Neural Network model of implementation is being done here for the segmentation procedure.

Image classification technique of CNN takes image as an input, processes it and subdivides it into various categories. Computers recognize an input image in the form of array of pixels and it solely relies on image resolution. It passes through various convolutional layers with filters and makes use of necessary functions to classify an object.

IV. System Architecture

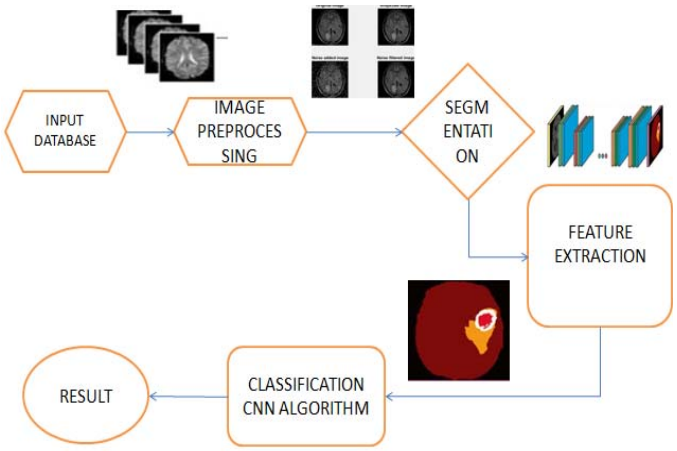


Fig. 1. System Architecture

Brain Tumor Segmentation using CNN is done mainly by using three modules.

[1]Image Preprocessing
[2]Image Segmentation
[3]Image Classification

V. Image Preprocessing

Digital image processing employs highly sophisticated algorithms for various virtual pictures. Digital image processing acquired many advantages when compared to analog photograph processing. Many algorithms are implemented at input statistics which avoids problem of noise distortion is shown in Fig. 1 to Fig. 8.

A. Image Intensification

Image intensification results in obtaining an image which is more suitable for future tasks. It includes noise removal, de-blurring, filtering. It needs to satisfy two conditions:
(i)The expected information must exist in the image.
(ii)The expected information must not be swamped by noise. It is of two types Qualitative and Quantitative.

Spatial Domain

Spatial Domain technique directly deals with the image pixels and manipulates to achieve the desired enhancement.

B. Frequency Domain

We process an image by changing the frequency components in an image without involving any convolutional operations.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:38:32 UTC from IEEE Xplore. Restrictions apply.

C. Filtering Technique

Filtration of an image was done for removing commotion, edge detection and increasing the contrast. Filtering can be of sequential and non-sequential. Sequential filters are often known as convolutional filters as they work with matrix operation employment. Non-Sequential filters perform in equalization to Median filtering.

D. Median Filtration Methodology

Median Filtration is a non-sequential technique for expelling commotion from images. It is broadly preferred as it obtains effectiveness in evacuating clamor without any disturbance to edges. This method is at most preferred in the removal of "pepper and salt" type of noise.

$$MSE = Sum(Sum((I(1,1) - K(1,1)).^2)) /(row * col) \quad (1)$$

$$PSNR = 10 * log10(255 * 255 / MSE) \quad (2)$$

MSE = Mean Square Error
PSNR = Peak Signal to Noise Ratio

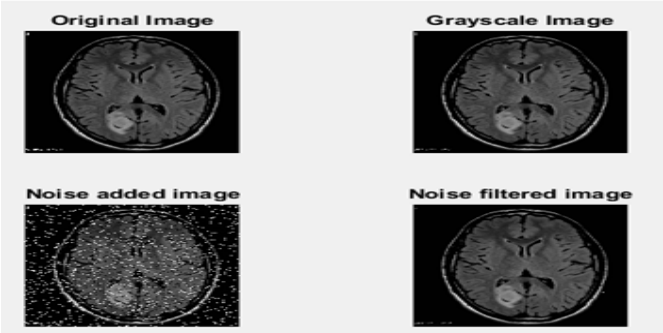


Fig. 2. Original to Noise Filtered Image

VI. IMAGE SEGMENTATION

The approach is generally useful in distinguishing of objects along with some pertinent data within that picture. The main intention of the procedure lies in simplification and rearranging an image into representational form which is progressively significant and helps in analyzing quickly. The procedure is accomplished by two ways.

(i)Discontinuity of power
(ii)Likeliness of force

The primary strategy involves segmenting a picture based on unexpected variation of color, intensity and texture. In latter one, based on regional division, an image is segmented in accordance with the group of formerly defined constraints.

A. Clustering Using K-Means

This technique implies in performing separation of items into groups which are comparable among each other and divergent to the items which are not a part of that cluster. It is a repetitive methodology highly effective in dividing an image into clusters of 'k'.

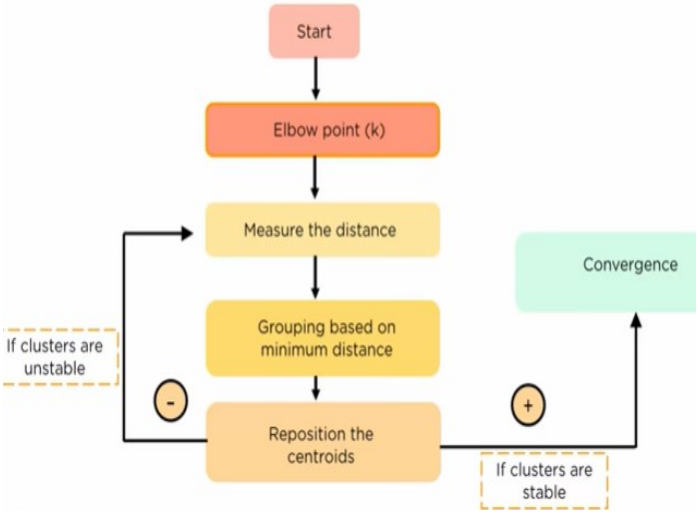


Fig. 3. Pictorial representation of Clustering Using K-Means

B. Procedure

Assume we have inputs x1, x2, x3, ….., and the value of k.

(i)	In the initial procedure, select k-irregular focuses in the form of centers for clusters known as centroid.
(ii)	The image pixels are aligned to clusters in such a way that the distance of the cluster center from pixel is reduced.
(iii)	Take the mean values of cluster pixels and refigure the environment
(iv)	Iterate the above two steps till we obtain convergence.

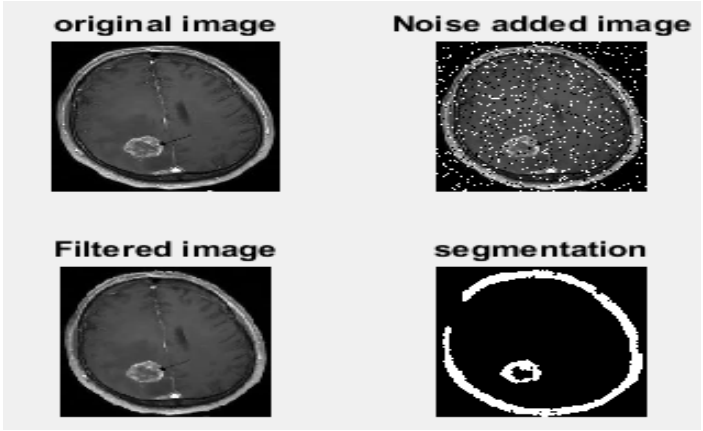


Fig. 4. Output of Image Segmentation

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:38:32 UTC from IEEE Xplore. Restrictions apply.

VII. IMAGE CLASSIFICATION

A. CNN Methodology

The most recognized and efficient algorithm in the area of multi median deep learning is CNN. Similar to the ones which are related to framework, it consists of data layer, covered layer and yield layer. They execute any one from the given methods (i)intricacy (ii)agglomeration and (iii)rectified sequential measure. Intricacy aligns given input images in a batch of sieves, which further performs photograph enhancement. Agglomeration enhances the output by compressing the network framework. Rectified Sequential measure helps in increasing speed and divides the training by delinerating non-positive values for zero without affecting any positive values. The three methods are iterated for all layers with obtaining distinguishing features. The wholly layer connector results a proportional vector of k, where k relates to category prediction number of network. The rearmost layer in the cnn lattice makes use of normalized exponential function in delivering classifying result.



Fig. 5. Segmented Part

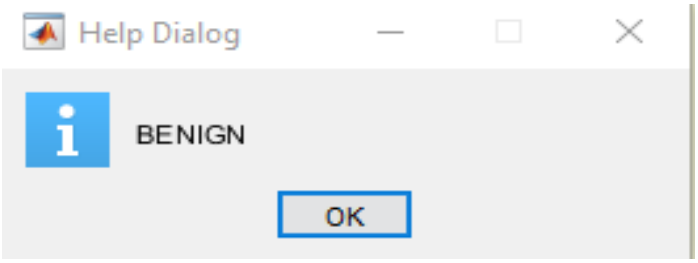


Fig. 6. Final Result

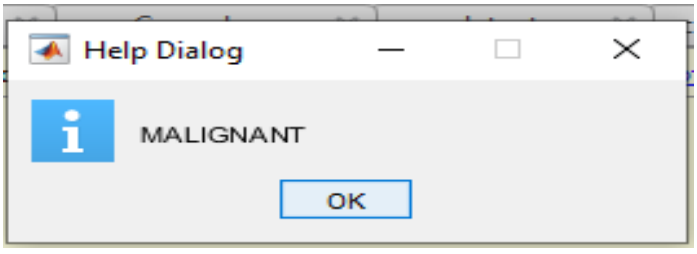


Fig. 7. Final Result

The below graph represents the accuracy of various techniques in detecting tumor when compared with CNN methodology.

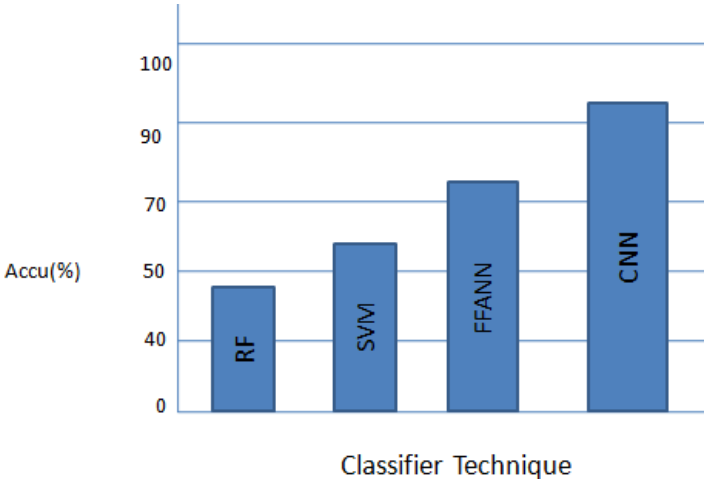


Fig. 8. CNN accuracy in comparison with other techniques

VIII. CONCLUSION AND FUTURE WORK

The Input databases has been taken out to the preprocessing and images has been filtered and noise has been removed. Image Segmentation has to be carried out by making use of cluster technique called K-means, then followed by Image Classification where the mri scanned images are classified using cnn algorithm to obtain more accuracy when compared to other algorithms and finally it shows the kind of tumor after performing classification. We can enhance work in future, where we can get the accurate tumor percentage and in detecting tumor to which the input data have been affected.

REFERENCES

[1] M. Prastaawa, "New segmentation in brain tumor framework on the basis of outlier detection", Med. Image Anal., vol. 18, no. 13, pp. 275-283, 2004.

[2] V. Rao, M. Sharifi, A. Jiswal, "Deep learning techniques in Brain tumor segmentation", MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pp. 76-59, 2016.

[3] A. Davy, Brain tumor segmentation based on deep neural networks MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS) (2016), pp. 1-5.

[4] A. Hamamci, Tumor-Cut: segmentation for brain tumors on contrast enhanced MR images for the usage in radio surgery applications IEEE Trans Med Imaging, 31 (3) (2013), pp. 790-809.

[5] Fiidon, L., Li, W., Garcia-Pereza-Herrara, L.C., Ekanayake, N., Oursaelin, S., Vercauterren, T.: Large Scalable multimodal convolutional networks in brain tumor segmentation. In: MICCAI, pp. 285–293 (2015).

[6] Raao, V., Sarabi, M. S., & Jaiiswal, A. (2016). Brain tumor segmentation using deep learning techniques MICCAI Multimodal Brain Tumor Segmentation Challenge (BraTS), pq. 56– 59.

[7] Ereira, S., Oliveira, A., Alves, V., & Silva, C. A. (2017). Basis hierarchical brain tumor segmentation for MRI using fully convolutional neural networks: A preliminary study. In 2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG) (pp. 1– 4).

[8] Milletari, F., Nava, N., Ahmadi, S.A.: V-Net: Entire convolutional neural networks for volumetric medical image segmentation. In: Fifth International Conference on 3D Vision (3DV) (2016)

[9] Ejaz, K., Rahim, M. S. M., Rehman, A., Chaudhary, H., Saba, T., and Ejaz, A., segmentation method in case of pathological brain tumor and accurate detection using MRI. Int. J. Adv. Comput. Sci. Appl. 9(8):394–401, 2018.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:38:32 UTC from IEEE Xplore. Restrictions apply.

[10] Reddy, B. V., Reddy, P. B., Kumar, P. S., and Reddy, S. S., Implementing an Approach to Brain MRI Image Preprocessing in Tumor Detection. International Journal of Research 1:725–731, 2015.

[11] Ism A, Direkoglu C, Sah M. review of MRI based brain tumor image segmentation on the basis of deep learning methods. In proceedings of 14th International Conference on Application of Fuzzy Systems and Soft Computing; August 29-30; Vienna, Austria: 2016

[12] Dong H, Yang H, Liu F, No Y, Guo Y. Automatic brain tumor detection and segmentation using U-net symbolized fully convolutional networks. MIUA. 2017;3:1–12.

[13] Jithina Jose, SujaCherukullapurathMana, B.Keerthi Samhitha, "An Efficient System to Predict and Analyze Stock Data using Hadoop Techniques", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2, July 2019

[14] Subhashini, R., Keerthi Samhitha, B., Mana, S.C., Jose, J,"Data Analytics To Study The Impact Of Firework Emission On Air Quality: A Case Study", AIP Conference Proceedings. 2019

[15] Samhitha, B.K., Mana, S.C., Jose, J., Mohith, M., Siva Chandhrahasa Reddy, L." An efficient implementation of a method to detect sybil attacks in vehicular ad hoc networks using received signal strength indicator", International Journal of Innovative Technology and Exploring Engineering, (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November, 2019

[16] Mana, S.C., Samhitha, B.K., Jose, J., Swaroop, M.V., Reddy, P.C.K., "Traffic violation detection using principal component analysis and viola jones algorithms", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, 8(3), September 2019

[17] Kanchana.R, Rajalakshmi.J, Mukesh.M.S, Nagarajan.V (FEB 2014), "Evolution of multi sensor image fusion using decimated biorthogonal wavelet transform and undecimated non-orthogonal wavelet transform", Int Journal of Electronic Communications Engineering Advanced Research, Vol.2, SP-1, Feb.2014, pp. 72-75

[18] Subhashini, R., Jeevitha, J.K., Samhitha, B.K." Application of data mining techniques to examine quality of water", Int J of Innovative Tech and Exploring Engg (IJITEE), ISSN: 2278-3075, 8(5S) 2019

[19] Surendran, R., Keerthi Samhitha, B.," Energy aware grid resource allocation by using a novel negotiation model", Journal of Theoretical and Applied Information Technology, 2014

[20] Jithina Jose, T Sasipraba , "Indoor air quality monitors using IOT sensors and LPWAN", Proceedings of IEEE: September 2019

[21] R.Vignesh , B.Keerthi Samhitha, C.M Suja, S.Divya, "Privacy Preserving Ride Sharing Scheme with Global Social Network for vehicles using Big Data, Indian J of Sci and Tech, 11(21) 2018

[22] Ramamoorthy, V., et al. " Examining and sensing of artificial knee with multi sensors networks", J of Advanced Research in Dynamical and Control Systems, 10(11) Pp. 115-120, 2018.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:38:32 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Analysing Android App Privacy Using Classification Algorithm

Ch. Chethan Sai	Ch. Sai Prakash	Jithina Jose	Suja Cherukullapurath Mana	B Keerthi Samhitha
Department of CSE	Department of CSE	School of Computing	School of Computing	School of Computing
Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology
chekuri.chethansai999@gmail.com	prakashnaidu073@gmail.com	jithinajose@gmail.com	cmsuja@gmail.com	samhitha711@gmail.com

ABSTRACT- The interface permits the client to scan for a subjective utility on the Play Store; the authorizations posting and the protection arrangement are then routinely recovered, on all events imaginable. The client has then the capability of choosing an interesting authorization, and a posting of pertinent sentences are separated with the guide of the privateer's inclusion and introduced to them, alongside a right depiction of the consent itself. Such an interface allows the client to rapidly assess the security-related dangers of an Android application, by utilizing featuring the pertinent segments of the privateer's inclusion and by introducing helpful data about shrewd authorizations. A novel procedure is proposed for the assessment of privateer's protection approaches with regards to Android applications. The gadget actualized widely facilitates the way toward understanding the security ramifications of placing in 1/3 birthday celebration applications and it has just been checked in a situation to feature troubling examples of uses. The gadget is created in light of expandability, and correspondingly inclines in the strategy can without trouble be worked in to broaden the unwavering quality and adequacy. Likewise, if your application handles non-open or delicate individual information, it would be ideal if you also allude to the extra necessities in the "Individual and Sensitive Information" territory underneath. These Google Play necessities are notwithstanding any prerequisites endorsed by method for material security or data assurance laws. It has been proposed that, an individual who needs to perform the establishment and utilize any 1/3 festival application doesn't perceive the significance and which methods for the consents mentioned by method for an application, and along these lines sincerely gives all the authorizations as a final product of which unsafe applications furthermore get set up and work their malevolent leisure activity in the rear of the scene.

I. INTRODUCTION

Versatile units of the scope of running frameworks have shown a lofty extend in the earlier decade, as needs are fundamental to make greater in the number and assortment of purposes that sudden spike in demand for cell phones. Cell phones are as of late utilized either in a reciprocal way to a processing gadget PC or even to substitute it. A closer shows up at the rationale of the utilization of PDAs unveil that they are now and again utilized for net perusing, long-range informal communication, and on-line banking. Likewise, they are utilized for portable explicit capacities, for example, SMS informing, read-time broadcasting of area, and pervasive access. As the capacities in cell cellphone usefulness grow (e.g., purposes for non-general wellbeing), cell phones come to be more noteworthy and extra engaging for a more extensive populace. Market information reviews uncover that the assortment of cell phone salary worldwide arrived at 208 million out of 2012.

This used to be a 38.3% stretch out contrasted with the deals in 2011. The improvement of cell phones and cell applications has brought about an essential exchange individuals' method for performing responsibilities in various variables of the day by day life, for example, making venture and directing social

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:37:09 UTC from IEEE Xplore. Restrictions apply.

correspondence. Cell phone capacities flaunt a rich range no longer just in a like manner consistently life activities anyway additionally for clients with additional specific needs. From computer games to mixed media applications, route frameworks, and wellbeing related applications, nowadays convenient cell application markets, for example, Google's Play Market, Apple's App Store, or Microsoft's Windows Store offer an enormous scope of capacities to clients with various requirements. The significant utility markets have been becoming consistently both in expressions of the capacities offered to the clients and the downloads did by method for the clients. The fast amplification in the acknowledgment of cell phones has prompted an expansion in their conceivable as an objective for pernicious exercises. This is common because the real clients' supply gets the right of section to form a scope of sorts of private data with the guide of the limit of cell applications. Therefore, a few applications in the market have been perceived as performing vindictive exercises. The unfurl of malevolent programming program is additionally impacted using the inclusion of the market suppliers. For example, Apple's App Store recently applies an inclusion that is dependent upon severe enrollment and friends gave advanced affirmation sooner than the dispatch of any application, in this way providing an assurance investigate the applications recorded in their utility stage usually. Others, for example, Google's Play Market, acquaint additional opportunities with engineers in transferring their one of a kind programming system to the market. The intentions are then dispensed with from the market if there should arise an occurrence of revealed noxious movement. Specifically, the Android working machine (in its present structure) licenses expulsion of the malignant utility from the gadget remotely. A comparable component of expulsion is likewise utilized using Apple's App Store. The inclusion of post-recognition expulsion of vindictive programming brings the need for early discovery of malware applications.

II. LITERATURE SURVEY

Composing review is the most crucial increment in the programming improvement process. Before making the framework, it is essential to make sense of the time factor, money related framework, and companions' quality. At the point when these issues are satisfied, at that factor, the following advance is to perceive out which working device and language can be used for building up the gadget. At the point when the product program engineers initiate creating the instrument, the developers need a bundle of open-air help. This guide can be purchased from senior engineers, from books or destinations. Before building the contraption, the above ideas are considered for developing the proposed structure. The hugest bit of the basic region considers and completely outlines all the basic requirements for endeavoring [3]. For each undertaking, a Literature review is the most basic stage in the programming improvement process. Before making the equipment and the related organizing it is crucial to decide and find out about the time factor, resource essential, work, economy, and boss quality. At the point when these issues are peaceful and considered, by then the following development is to make sense of about the item specs in the man or lady device, for instance, what sort of working system the action would require, and what is generally the quintessential writing computer programs are wished to continue with the subsequent advance, for instance, developing the gadgets, and the related errands[4]**.**

[1]Versatile malware has been developing in scale and intricacy prodded through the unabated take-up of cell phones around the world. Android is quickly transforming into the most renowned cell stage resulting in sharp reach out in malware focused on the stage. Moreover, Android malware is advancing quickly to avoid location by methods for basic mark based checking. Despite current location quantifies in place, auspicious disclosure of new malware is in any case a basic issue. This calls for novel ways to deal with alleviating the creating peril of zero-day Android malware. Consequently, the creators fortify and break down proactive AI methods dependent on Bayesian arrangement planned for revealing obscure Android malware with the guide of static examination. The investigation, which depends on a monstrous malware test set of lion's shares of the present families, exhibits recognition capacities with exorbitant precision. Exact results and near investigation are presented providing advantageous observation toward the improvement of successful static-systematic Bayesian grouping based choices for recognizing obscure Android malware.

[2]The additional exceptional blast of the Android stage in the past hardly any years has made it a worthwhile objective of vindictive programming (application) designers. There are several circumstances of malware applications that transport

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:37:09 UTC from IEEE Xplore. Restrictions apply.

premium value SMS messages, melody clients' non-open information, or applications that, regardless of whether never again described as malware, conduct faulty moves influencing the client's protection or costing them cash. Right now, research the achievability of the utilization of both the authorizations an application demands, the classification of the application, and what consents are mentioned using various applications in the indistinguishable class to higher education clients whether the perils of introducing an application is comparable with its foreseen advantage. Existing methods think about on consideration exclusively the dangers of the consents mentioned by an application and discard both the points of interest and what authorizations are mentioned by different applications, therefore having a limited impact. The endless hazard flags are prompted and assess them utilizing two datasets, one comprises of 158,062 Android applications from the Android Market, and another comprises of 121 noxious applications. The adequacy of our motivation is shown across the board records examination.

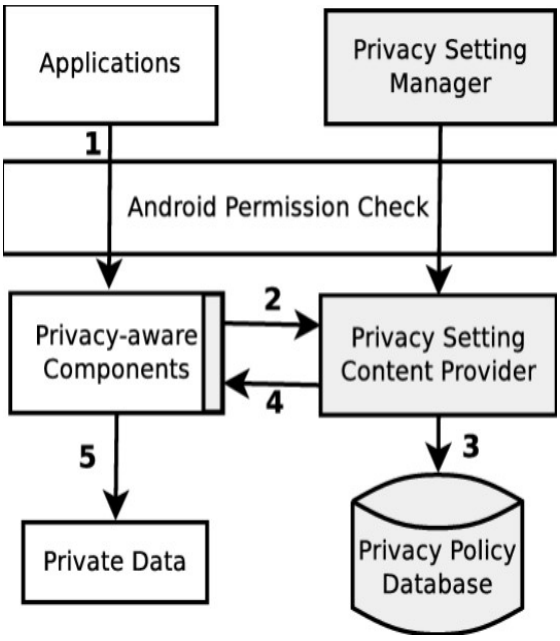


Fig. 1. System Architecture.

Recognize the rundown of outsider set up applications. Concentrate the total posting of consents of every application. Distinguish android: assurance phase of every consent, for example, Typical or Dangerous of each application. Take the android application consents dataset. To recognize the perilous applications, watch order calculations. Note the precision of spontaneous mail grouping given by it and the time required for execution. Results as precision among particular perilous and normal applications Classifiers are examined.

This is done in five different steps as follow

- LOGIN MODULE
- ENLISTMENT MODULE
- APPLICATION DOWNLOAD
- ANDROID PERMISSION CHECK
- SECURITY POLICY DATABASE

1. LOGIN MODULE

This is the main endeavor that opens when the customer opens the site. The client needs to supply a right contact run and a secret phrase, which the client enters while enlisting, to login into the site. On the off chance that records provided by a method for the client suits with the realities in the database work area then the client effectively login into the site else message of login fizzled is shown and the client needs to return the right data. A connection to the register action is moreover outfitted for enrollment of new clients.

2. ENLISTMENT MODULE

Another individual who wishes to get passage to the application needs to enroll first sooner than login. By tapping on the register button in login action, the register diversion gets open. Another client enrolls by utilizing entering a complete name, secret phrase, and contact number. A client needs to enter secret word again in insist secret phrase textbox for affirmation. At the point when the client enters the data in all text boxes, on the snap-on of the register button, the data is moved to the database, and the individual is coordinated to login undertaking once more. An enrolled customer at that point wishes to log in to get admission to the application. Approvals are applied to all the textboxes for the perfect working of the application. Like realities in each textbox is have to

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:37:09 UTC from IEEE Xplore. Restrictions apply.

that is each textbox, it may be of name, contact, secret key or affirm

secret phrase, will now not be vacant while enrolling. On the off chance that any such textbox is the unfilled application will supply the message of measurements is have to in each textbox. Additionally, the information in secret word and confirm secret key fields need to fit for beneficial enrollment. Another approval is contact assortment should be a genuine one that is of 10 digits. On the off chance that any such approval is abused, at that point enlistment will be ineffective and afterward, the client needs to enroll once more. The message that the application will show when one of the fields is vacant. If every such datum is the right client will be coordinated to a login leisure activity for login into the application.

3. APPLICATION DOWNLOAD

Right now, applications are as of now download in the cell phone using play store. After establishment just find the malware current in an assortment of applications.

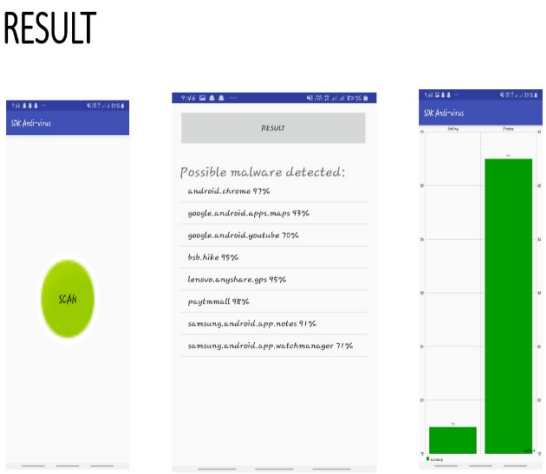
4. ANDROID PERMISSION CHECK

A client who wants to the establishment and utilizes any third festival application doesn't catch the worth and that methods for the consents mentioned through an application, and accordingly in actuality offers all the authorizations because of which unsafe applications moreover get introduced and play out their pernicious exercise at the rear of the scene.

5. SECURITY POLICY DATABASE

A novel systems is proposed for the assessment of security strategies with regards to Android applications. The gadget considerably facilitates the methodology of handling the protection ramifications on introducing third festival applications and it has just been checked in a situation to highlight troubling.

IV. RESULT AND CONCLUSION



CONCLUSION

Right now, the application handles non-open or delicate individual information, it would be ideal if you furthermore allude to the extra prerequisites in the "Individual and Sensitive Information" region beneath. These Google Play prerequisites are notwithstanding any necessities endorsed with the guide of significant security or measurement insurance laws. An individual who wishes to the establishment and utilizes any 1/3 gathering application doesn't secure the significance and which methods for the consents mentioned by an application, and in this way truly offers all the authorizations because of which unsafe applications also get snared and work their malevolent exercise in the rear of the scene.

REFERENCES

[1] Yerima, S. Y., Sezer, S., & McWilliams, G. (2014). Analysis of Bayesian classification-based approaches for Android malware detection. IET Information Security, 8(1), 25-36.

[2] Sarma, B. P., Li, N., Gates, C., Potharaju, R., Nita-Rotaru, C., & Molloy, I. (2012, June). Android permissions: a perspective combining risks and benefits. In Proceedings of the 17th ACM symposium on Access Control Models and Technologies (pp. 13-22). ACM.

[3] Chia, P. H., Yamamoto, Y., & Asokan, N. (2012, April). Is this app safe?: a large scale study on application permissions and risk signals. In Proceedings of the 21st international conference on World Wide Web (pp. 311-320). ACM.

[4] Yang, Z., Yang, M., Zhang, Y., Gu, G., Ning, P., & Wang, X. S. (2013, NovemberAppintent: Analyzing sensitive data transmission in android for privacy

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:37:09 UTC from IEEE Xplore. Restrictions apply.

leakage detection. In Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security (pp. 1043-1054). ACM.

[5] Zhang, W., Li, X., Xiong, N., & Vasilakos, A. V. (2016). Android platform- based individual privacy information protection system. Personal and Ubiquitous Computing, 20(6), 875-884.

[6] Felt, A. P., Chin, E., Hanna, S., Song, D., & Wagner, D. (2011, October). Android permissions demystified. In Proceedings of the 18th ACM conference on Computer and communications security (pp. 627-638). ACM.

[7] J. Jose and T. Sasipraba, *"Indoor air quality monitors using IOT sensors and LPWAN,"* 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 633-637

[8] "*An Efficient Implimentation of Methods to Detect Sybil Attacks in Vehicular Ad hoc Networks Using Received Signal Strength Indicators* ",B.Keerthi Samhitha , SujaCherukullapurathMana **, Jithina Jose**, M. Mohith, L. Siva Chandhrahasa Reddy International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November, 2019

[9] "*An Efficient System to Predict and Analyze Stock Data using Hadoop Techniques*" , **Jithina Jose**, SujaCherukullapurathMana, B.Keerthi Samhitha International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2, July 2019

[10] "*Data Analytics To Study The Impact Of Firework Emission On Air Quality: A Case Study* "Dr.R .Subhashini,B.Keerthi Samhitha, SujaCherukullapurathMana ,**Jithina Jose** Proceeding of 3rd Intenational Conference on Trends in Electronics and Informatics (ICOEI 2019) SCAD College of Emgineering ,Thirunelveli , April 23- 25 2019

[11] *"Traffic violation detection using principal component analysis and viola jones algorithms"*, SujaCherukullapurathMana, **Jithina Jose,** B.Keerthi Samhitha International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, September 2019.

[12] W. Chen and Y. Gao, "A Minutiae-Based Fingerprint Matching Algorithm Using Phase Correlation," 9th Biennial Conference of the Australian PatternRecognition Society on Digital Image Computing Techniques and Applications (DICTA 2007), Glenelg, Australia, 2007, pp. 233-238.doi:10.1109/DICTA.2007.4426801

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:37:09 UTC from IEEE Xplore. Restrictions apply.

An Integrated and Cohesive Methodology – Smart Manufacturing Service

L.Mary Gladence [1], A.V.E.Jayakiran [2], A.Indraneel [3], V. Maria Anu[4]
[1].[4] School of Computing, Sathyabama Institute of Science and Technology, Chennai, India.
[2][3] Student, CSE Department, Sathyabama Institute of Science and Technology, Chennai, India.

***Abstract*—Smart manufacturing systems on manufacturing devices connected to manufacturing equipment are efficiently implemented. This is a challenge and is extremely beneficial in encouraging the development of a smart plant so that all manufacturing equipment are connected, plug-and-play and accessible via the network. Cloud production is a significant or to offer highly customized services to manufacturing companies on the Internet. This paper explores the promotion for successful services-oriented undertakings of cloud manufacturing systems. In order to deal with this issue, this program provides a cloud-based pluggable production process through the use of cloud computing, cutting edge computing and restful web services. The production services can be produced in the form of a pluggable application module (PAM) by using a two-layer service mechanism.The framework proposed enables engineers to deploy selected cloud PAMs to effectively target edge devices and to run and manage the plugged PAMs over the remote cloud platform using Web-based GUIs in order to facilitate smart management on target manufacturing equipment. Enhanced manufacturing-network visibility, knowledge exchange, alignment and flexibility of production processes and informed decision-taking-are key elements of the new program, such that development output is plan-based.**

Keywords: Smart manufacturing services; Edge computing; Pluggable manufacturing service scheme.

I INTRODUCTION

Through the success of assembly developments, hardware quality is increased more and more. Therefore, making clever capabilities for hardware that guarantees its reliability and creating consistency, for example FDC (Fault Detection and Classification), fabricating accuracy guessing, RUL (Remaining Useful Life), etc., is an issue for the company. The application of these skills to hardware fits the trend for cleverer computers and frameworks [1]. Cloud processing (CMFG)[2] has recently become a hot topic for research, both in the academic and industrial sectors. CMFG's concept is to show and virtualize appropriate manufacturing propertiesCloud administration (i.e. manufacturing administrations) will be able to use systems on-demand (including item structure, replication, generation, check, support) and any other company in the item life cycle in cloud administrations (equipment and programming). Because of the inexhaustible assets of distributed computing, these assembling cloud administrations can bolster plant wide or cross-plant fabricating exercises [3].With regard to cloud dormancy, it is increasingly sufficient for some assembly administrations that use edge processing to provide their functions with certain time-based equipment. For example, the machining business has proposed the idea of Smart Machine Box (SMB)[4], that uses a rim-setting gadget, a short range of information and capacity data, a lack of location and a alert, correspondence, items and machine visuals, as an interface with a machine device to provide power.One more concept is the AMCOT (Advanced Manufacturing Cloud of Things) [5] genius assembly stage CPAs. Every CPA is made on a borderlineto supply and prepare information and algorithm modules and a correspondence in a shrewd article that enables AMCOT to be thought of in digital physical associations. Usually, a bit of generation gear requires layered pieces. In the sun driven cell production phase, for example, the PECVD hardware involves parts at module level, including the vacuum module, temperature module, RF module, and components of a gadget such as pressurizing mechanism, throttle valves, vacuum pressurizing. Moreover, an assembling manufacturing plant has a lot of hardware. Thusly, actualizing shrewd assembling administrations on

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.

processing plant wide edge gadgets associated with gear effectively so those assembling administrations are pluggable, attachment and-play, and reasonable through the system is a difficult undertaking and is exceptionally advantageous for encouraging understanding a savvy manufacturing plant. Conventional methodologies, for example, pre-introduced programming or downloading Apps on request tense gadgets, can't accomplish the previously mentioned vision. No writing has tended to this issue.

Building robotization frameworks utilize open correspondence principles and interfaces that can incorporate a wide decision of various structure control teaches, for example, warming, ventilation, cooling, lights and blinds, security highlights, and hardware. Prior structures, be that as it may, don't for the most part have these frameworks. By and large, each sort of establishment offers explicit administrations: Heating Ventilation and Air Conditioning (HVAC) controls atmosphere administrations, cameras and sensors offer security administrations, and so forth. Just when a vitality the executive framework is planned, the various subsystems are connected, however just associated through the structure's vitality the executive framework. The vitality the board administrations are incorporated in particular programming. Various makers with various administrations think that it hard to coordinate new administrations and functionalities, for -the two clients and support professionals. Computerized structures coordinate programming for control and information securing with modern conventions and interfaces. What's more, coordinating new administrations into this kind of arrangement isn't simple. It relies upon the introduced programming's advancement. These mechanical advancements additionally offer cloud-associated arrangements and wise administrations for vitality the executives. These administrations are likewise evolved in incorporated PC frameworks. The information are transmitted to these frameworks or to the cloud, where they are dissected.

This paper proposes to utilize the edge and mist processing ideal models sent in Internet of Things (IoT) advancements with two fundamental points: 1. To encourage the coordination of new wise and interoperable administrations in robotized and non-mechanized structures (reconciliation). 2. To permit the circulation of savvy benefits between the entirety of the structure's subsystems (interoperability). Through this proposition, the connection between a structure's subsystems is encouraged. It likewise advances the formation of new savvy administrations (e.g., new dispersed brilliant control calculations, utilizing information caught by power the board to distinguish human action, catching example acknowledgment of gadget association, computing sustainable force expectation, utilizing power information in security administrations, and so forth.). Right now, planned a middleware design with two principle layers dependent on installed gadgets, IoT correspondence conventions and equipment backing to create computerized reasoning calculations. To accomplish this, we added two calculated levels to the structure's offices: edge and mist hub. Each level has various types of gadgets and functionalities. We conveyed and actualized a layer-based middleware design to test the model.

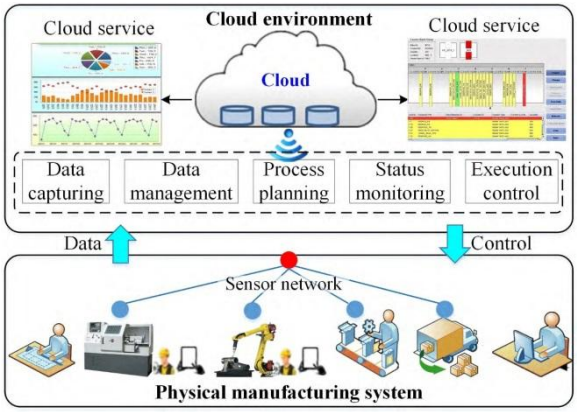


Fig 1 Cloud Environment

Cloud fabricating is developing as a key empowering agent for assembling organizations to convey profoundly customizable administrations over the Internet. Executing wise assembling administrations on manufacturing plant wide edge gadgets associated with creation hardware effectively[21-26]. It is sensible through the system and is additionally testing task which is profoundly advantageous for encouraging understanding a savvy manufacturing plant. This work means to explore how cloud producing frameworks can encourage compelling help situated business. This work

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.

proposes a cloud-based pluggable assembling administration plot by utilizing distributed computing and soothing Web Service.

II RELATED WORK

The system was as of late labeled at two closures called the "edge" and "center" finishes to pinpoint where the handling happens [1]. The edge end is near the information sources and clients, and the center end comprises of the cloud servers. Right now, edge figuring worldview [2] pushes calculations to the edge of the IoT arrange [3], to lessen information preparing latencies and the quantity of information sent to the cloud. A cloud-based backend can serve handling demands that are less time-delicate or of which results are not required by the source gadgets themselves (e.g., huge information examination on the condition of the IoT organize).

As far as edge processing asset provisioning, the continuous Horizon 2020 RECAP venture [4] proposes a coordinated cloud–edge–mist design planned for fathoming application position, foundation the board and limit provisioning. Cloud/Edge framework observing is advanced with application, foundation and outstanding task at hand models, which are thus bolstered into an improvement framework that coordinates applications and constantly designs the framework.

Xiang et al. [5] developed a mechanical technology software work system. Nevertheless, information obtained from the machinery and sensors is transmitted to the cloud and the desired yields re-sent to the desired gadgets. Besides, monstrous information requires high data transfer capacity, which is costly [6]. With arrange inaccessibility, these issues become boundaries to applications required defer limitations [7]Haze processing and the edge formulation are the answers to above problems, which include transferring some cloud registry, storage and device management capabilities to the edge [8]. By mist and border recording, any information from informed machines or sensors should not be placed away or made ready on a computer, but rather anxious gadgets. (e.g., haze switches or keen entryway/IIhub [9], and so forth.), which satisfies the necessities like low idleness, ongoing reaction, decrease of system traffic, and so on [10].Furthermore, the mist is also able to track and lower the loss of assets[11]. As a new global view of production, haze registration and cutting edge appearance are likely to increase efficiencies and build innovative management[12] to allow the development of skillful manufacturing to be speeded up. A reference design is proposed to guide progress of the keen assembly, in this way, for the shrewd assembly of anxious production, mist registration and distributed computing. In the computerized twin shop floor, the reference design will be applied to increase different levels of shop floor insight. The examination directed by ***Xu et al. [5]*** proposes an utility-mindful asset distribution technique for edge registering, called Zenith. With Zenith, specialist co-ops can set up asset imparting agreements to edge framework suppliers permitting an inactivity mindful asset provisioning calculation to plan edge undertakings in a way that meets their inertness necessities. Edge Node Resource Management (ENORM) [13] is a structure that oversees edge/mist hub assets, empowering auto-scaling edge hubs by observing application needs. Explicit applications can be organized through static need task. Provisioning and auto-scaling components are moderately straightforward executions dependent on direct hunt. The edge cloud worldview is likewise practical when the sources themselves are portable [14]. Keen computational offloading of cell phones towards edge hubs, explicitly at the edge of radio access systems, was considered by ***Chen et al. [8].*** Right now, creators gave task offloading calculations planning the appropriated calculation offloading choice as a multi-client calculation offloading game. In a similar line, ***Wang et al. [9]*** researched the issue of together planning offloaded undertakings to a huge number of edge hubs and propose presenting confirmation control at the edge level alongside a two-stage booking approach, improving offloading execution contrasted with conventional nearest edge determination draws near.

III EXISTING SYSTEM

The current framework comprises of all the assembling assets associated with different assembling exercises. Material comprises of crude

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.

material, partsHead assets may be cut off as "Human-Machine-Material-Environment" in a shop floor or in the manufacture line. Human beings apply to staff and administrators. The system involves machining equipment (e.g., machines, robots, etc.) moving equipment (e.g. components, programmed guided vehicles, etc.) and segments, semi-completed items, completed items, and so forth. In the current framework, is liable for medium speed and idleness information examination for assembling data reconciliation and the board. The parts (e.g., arrange switches, territorial servers, organize switches, and so on.) of edge or center system, and undertaking class the board frameworks could be utilized as the foundations for mist processing. The haze the board hub can know about all the mist hubs in a mist arrange. The mist hub speaks with the cloud server farm of the outer system through the mist the board hub. In haze layer, edge association gadgets and edge figuring gadgets speaks with the haze door through a solitary bounce low-dormancy interface. The internal mist hubs speak with the edge layer through the haze portal. The primary capacity of the internal haze hub is to perform information handling (information separating, pressure, stockpiling, encryption, and so forth.) on the information from the edge layer and preparing administration demands. Information created by different assembling assets, which might be geologically circulated, are encountering dangerous development. Theseinformation are passed on over the system to cloud, where information preparing is completed. The expanding volume and speed of information require high data transmission, which is over the top expensive. At the point when the system clog is not kidding, a few information might be misfortune.

IVPROPOSED SYSTEM

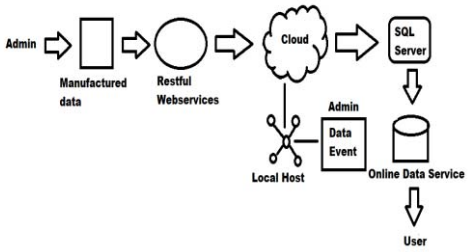


Fig 2 Overview of the Proposed System

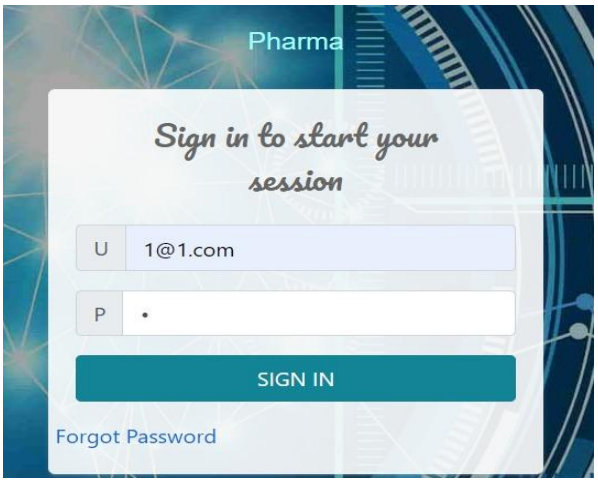
The proposed framework (Fig 2) manages multisource, multiscale and multivariate information. Multisource information: various data Sources are connected with the range of knowledge, IoT, Controlling Frameworks, Cyber-Physics Systems (CPS), Enterprise Information Systems (EIS) and others (e.g., computers, production lines, inventories, and so forth) assembling resources. Data multi-scale: Different kinds of information from different activities and procedures are included, which range over numerous degrees of the mechanical mechanization pecking order, from low level machines, to creation lines and store network systems, up to the entire processing plant advanced world. Multivariate information: The information display various factors, through which it tends to be considered and prepared, for example, transient perspectives that influence the determination of the fitting assortment components or the objective assets that this information is alluded as for the choices that should be made. In the proposed framework, the utilization case driven adaptable total of brilliant assembling information: shrewd assembling involves increasingly exact and complete control on circumstance mindfulness for the leaders and the partners in the business to empower them evaluate the effect of disturbances on the present creation and store network abilities.

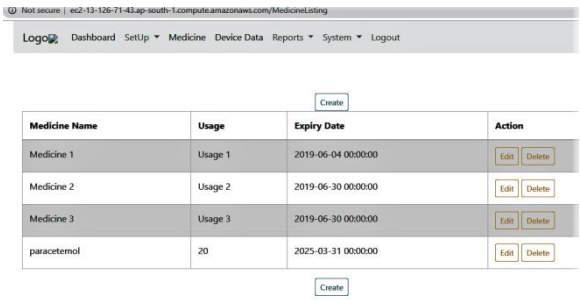
Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.

V. RESULTS AND DISCUSSION

The software for multi-smart growth. The underlying layer is the intelligent devices that provide edge computing and are the sources of data. The medium-layer is the fog computing data storage network. The cloud is the highest layer, where large data is processed and analyzed. By processing, storage and networking in close-end nodes, edge processing and fog computing decreases cloud data and the probability of service interruption, which ensures intelligent production system robustness. Cloud computing is more successful in meeting the demands of intelligent manufacturing applications.

Cloud will launch instances and provide the link. So as we can use the link in online services.





VI. CONCLUSION

For assembling organizations to remain serious, delivering new PSS using disseminated assets dependent on a pay-more only as costs arise model is a compelling methodology. There are as yet various specialized difficulties in changing the present assembling frameworks into administration arranged assembling models under a haze fabricating condition. The proposed cloud-based framework can encourage administration situated assembling business in a decentralized and dynamic assembling condition. The significance of information driven help organization in improving the dynamic procedures in brilliant assembling biological systems has been featured.

REFERENCES:

[1] F. Tao and Q. Qi, ``New IT driven service-oriented smart manufacturing: Framework and characteristics," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 1, pp. 81_91, Jan. 2019.

[2] C. Zhu, V. C. M. Leung, L. Shu, and E. C.-H. Ngai, ``Green Internet of Things for smart world," *IEEE Access*, vol. 3, pp. 2151_2162,Nov. 2015.

[3] L. Da Xu, W. He, and *Ind. Informat.*, vol. 10, no. 4, pp. 2233_2243, Nov. 2014.

[4] F. Tao, Q. Qi, A. Liu, and A. Kusiak, ``Data-driven smart manufacturing," *J. Manuf. Syst.*, vol. 48, pp. 157_169, Jul. 2018.

[5] J. Pizo« and J. Lipski, ``Perspectives for fog computing in manufacturing," *Appl. Comput. Sci.*, vol. 12, no. 3, pp. 37_46, 2016.

[6] Y. Wei and M. B. Blake, ``Service-oriented computing and cloud computing: Challenges and opportunities," *IEEE Internet Comput.*, vol. 14, no. 6, pp. 72_75, Nov. 2010.

[7] Q. Zhang, L. Cheng, and R. Boutaba, ``Cloud computing: State-of-the-art and research challenges," *J. Internet Services Appl.*, vol. 1, no. 1, pp. 7_18, May 2010.

[8] F. Tao, L. Zhang, V. C. Venkatesh, Y. Luo, and Y. Cheng, ``Cloud manufacturing: A computing and service-oriented manufacturingmodel," *Proc. Inst. Mech. Eng., B, J. Eng. Manuf.*, vol. 225, no. 10, pp. 1969_1976, 2011.

[9] F. Tao, L. Zhang, Y. Liu, Y. Cheng, L. Wang, and X. Xu, ``Manufacturing service management in cloud manufacturing: Overviewand future research directions," *J. Manuf. Sci. Eng.*, vol. 137, no. 4, p. 040912, 2015.

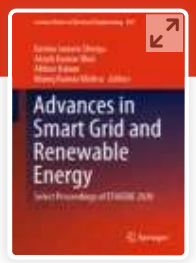
[10] Y. Liu and X. Xu, ``Industry 4.0 and cloud manufacturing: A comparative analysis," *J. Manuf. Sci. Eng.*, vol. 139, no. 3, p. 034701, Mar. 2017.

[11] A. F. da Silva, R. L. Ohta, M. N. D. Santos, and A. P. D. Binotto, ``A cloud-based architecture for the Internet of Things targeting industrial devices remote monitoring and control," *IFAC_PapersOnLine*, vol. 49,no. 30, pp. 108_113, 2016.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.

[12] G. Xiong, T. Ji, X. Zhang, F. Zhu, and W. Liu, ``Cloud operating system for industrial application,'' in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat.*, Nov. 2016, pp. 43_48.
[13] L. Shu, Y. Zhang, Z. Zhou, M. Hauswirth, Z. Yu, and G. Hynes, ``Transmitting and gathering streaming data in wireless multimedia sensor networks within expected network lifetime,'' *Mobile Netw. Appl.*, vol. 13, nos. 3_4,pp. 306_322, 2008.
[14] L. Shu, Y. Zhang, L. T. Yang, Y. Wang, M. Hauswirth, and N. Xiong, ``TPGF: Geographic routing in wireless multimedia sensor networks,''*Telecommun. Syst.*, vol. 44, no. 1, pp. 79_95, Jun. 2010.
[15] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, ``Fog computing and its role in the Internet of Things,'' in *Proc. 1st Ed. MCC Workshop Mobile Cloud Comput.*, 2012, pp. 13_16.
[16] F. Tao, J. Cheng, and Q. Qi, ``IIHub: An industrial Internet-of- Things hub toward smart manufacturing based on cyber-physical system,''*IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 2271_2280, May 2018.
[17] D.-J. Deng, S.-Y. Lien, C.-C. Lin, S.-C. Hung, and W.-B. Chen, ``Latency control in software-de_ned mobile-edge vehicular networking,'' *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 87_93, Aug. 2017.
[18] I. Stojmenovic and S. Wen, ``The fog computing paradigm: Scenarios and security issues,'' in *Proc. IEEE Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, Sep. 2014, pp. 1_8.
[19] F. Tao, Y. Cheng, L. D. Xu, L. Zhang, and B. H. Li, ``CCIoTCMfg: Cloud computing and Internet of Things-based cloud manufacturingservice system,'' *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1435_1442, May 2014.
[20] H. Lee, ``Framework and development of fault detection classification using IoT device and cloud environment,'' *J. Manuf. Syst.*, vol. 43,pp. 257_270, Apr. 2017.
[21] L.Mary Gladence,T.Ravi,Y.Mistica Dhas"An Enhanced Method for Disease Prediction using Ordinal Classification-APUOC"in Journal of Pure and Applied MicroBiology ISSN:0973-7510 November2015 Vol.9 Special Edition Nov 2015
[22] Subhashini, R., Mary Subaja Christo, G. Parthasarathy, and J. Jeya Rathinam. "Crowd Sourcing Application for Chennai Flood 2015." In International Conference on Emerging Current Trends in Computing and Expert Technology, pp. 1382-1388. Springer, Cham, 2019.
[23] Revathy M, Ramya S, Sathiyavathi R, Bharathi B, Anu VM. Automation of street light for smart city. In2017 International Conference on Communication and Signal Processing (ICCSP) 2017 Apr 6 (pp. 0918-0922). IEEE.
[24] Dr.R. Subhashini, R. Sethuraman, V. Milani, "Reinforcing Telemedicine Through an Interactive Voice Response Service for Rural Indians", International Journal of Engineering and Technology, ISSN: 0975-4024, Vol 7 No 1, Feb-Mar 2015
[25] S.Shanmuga Priya, A.Valarmathi, M.Rizwana, L.Mary Gladence, "Enhanced Mutual Authentication System in Mobile Cloud Environments" in International Journal of Engineering and Technology ISSN: 2227-524X, Vol.7 No.3.34 2018,Page.No 192-197
[26] Gladence, L. Mary, V. Maria Anu, R. Rathna, and E. Brumancia. "Recommender system for home automation using IoT and artificial intelligence." Journal of Ambient Intelligence and Humanized Computing (2020): 1-9.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:34:20 UTC from IEEE Xplore. Restrictions apply.



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Smart Grid and Renewable Energy** pp 659–666

An Integrated & Robust User Authentication Framework Based on Gait Smart Devices

V. Harsha Vardhan ✉, V. Sri Dileep Kumar, G. Nagarajan, T. Prem Jacob & A. Pravin

Conference paper | First Online: 05 January 2021

301 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

Stride acknowledgment is the mapping of a step grouping to a character known to the framework. Walk verification alludes to the issue of distinguishing whether a given stride arrangement has a place with the asserted personality. A common stride confirmation framework begins with a component portrayal, for example, a step layout, which separates its highlights, and changes are applied in the end to get a discriminant set. Every confirmation in writing utilizes Euclidean separation

as an edge to check the limit between a genuine subject and an impostor. This article proposes a technique that utilizes the Bayes' classifier instead of the Euclidean separation. The proposed system is applied to layout-based step to be included in portrayals. This is a novel structure for client confirmation system, dependent on examination by human intervention in association with smartphone sensors information. Being client friendly and easy to access, human step conduct investigated from smartphone sensors information is helpful and easy to use for client validation.

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-981-15-7511-2_67
- Chapter length: 8 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Softcover Book EUR 279.99

Hardcover Book EUR 279.99

Learn about institutional subscriptions

References

1. M. Shoaib, S. Bosch, O.D. Incel, H. Scholten, P.J.M. Havinga, Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors **16**, 426 2016 [CrossRef] [PubMed]

2. E. Miluzzo, C.T. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu, A.T. Campbell, Darwin phones: the evolution of sensing and inference on mobile phones. Darwin **14**, 5–20 (2010)

3. J. Kwapisz, G. Weiss, S. Moore, Activity recognition using smartphone accelerometers. ACM SIGKDD Explor. **12**, 74–82 2011 [CrossRef]

4. M. Mun, P. Boda, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. Hansen, E. Howard, R. West, PEIR, the personal environmental impact report, as a platform for participatory sensing systems research, In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services, Kraków, Poland, 22–25 June 2009*, pp. 55–68.

5. M. Shoaib, S. Bosch, H. Scholten, P.J.M. Havinga, O.D. Incel, Towards detection of bad habits by fusing smartphone and smart watch sensors, in *Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops,*

St. Louis, MO, USA, 23–27 March 2015, pp. 591–596.

6. D. Shukla, R. Kumar, A. Serwadda, V.V. Phoha, Beware, your hands reveal your secrets! In *Proceedings of the CCS-ACM Conference on Computer and Communications Security, Scottsdale, AZ, USA 3–7 November 2014*, pp. 904–917

7. A.J. Aviv, K. Gibson, E. Mossop, M. Blaze, J.M. Smith, Smudge attacks on smartphone touch screens, in *Proceedings of the 4th USENIX Conference on Offensive Technologies, Washington, DC, USA, 11–13 August 2010*, pp. 1–7

8. B. Draffin, J. Zhu, J. Zhang, KeySens: passive user authentication through micro-behavior modeling of soft keyboard interaction. Mob. Comput. Appl. Serv. **130**, 184–201 (2014)

9. A. Alzubaidi, J. Kalita, Authentication of smartphone users using behavioral biometrics. IEEE Commun. Surv. Tutor. **18**, 1998–2026 (2016) [CrossRef]

10. D. Gafurov, A survey of biometric gait recognition: approaches, security and challenges, in *Proceedings of the Norwegian*

Informatics Conference, Oslo, Norway 19–21 November 2007, pp. 19–21

11. D. Gafurov, E. Snekkenes, P. Bours, Spoof attacks on gait authentication system. IEEE Trans. Inf. Forensics Secur. **2**(3), 491–502 (2007)

12. A. Hadid, M. Ghahramani, V. Kellokumpu, M. Pietikäinen, J. Bustard, M. Nixon, Can gait biometrics be spoofed? in *Proceedings of 21st International Conference Pattern Recognition (ICPR), November 2012*, pp. 3280–3283

13. It's the era of the smartwatch: IDC says device to rule nearly half of wearables by 2022. https://economictimes.indiatimes.com/magazines/panache/its-the-eraof-the-smartwatch-idc-says-device-to-rule-nearly-half-of-wearables-by-2022/articleshow/65810524.cms. Accessed 1 Mar 2018 (2018)

14. G. Cola, M. Avvenuti, F. Musso, A. Vecchio, `Gait-based authentication using a wrist-worn device, in *Proceedings of 13th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MOBIQUITOUS), New York, NY, USA, Nov 2016*, pp. 208–217

15. S. Terada, Y. Enomoto, D. Hanawa, K. Oguchi, Performance of gait authentication using an

acceleration sensor, in *Proceedings of 34th International Conference on Telecommunications and Signal Processing (TSP), Aug 2011*, pp. 34–36

16. A.H. Johnston, G.M. Weiss, Smartwatch-based biometric gait recognition, in *Proceedings of IEEE 7th International Conference on Biometrics Theory, Applications, Systems (BTAS), Sept 2015*, pp. 16

17. A. Mahfouz, T.M. Mahmoud, A.S. Eldin, Poster: a behavioral biometric authentication framework on smartphones, in *Proceedings of ACM Asia Conference Computers and Communications Security (ASIA CCS), New York, NY, USA, Apr 2017*, pp. 923–925. https://doi.acm.org.proxy.library.stonybrook.edu/10.1145/3052973.3055160

18. R. Christina, G.L. Shajan, B. Ankayarkanni, CART—A statistical model for predicting QoE using machine learning in smartphones, in *IOP Conference Series: Materials Science and Engineering*, vol. 590, no. 1 (IOP Publishing 2019), p. 012001.

19. T. Rajalakshmi, R. I. Minu, Improving relevance feedback for content based medical image retrieval, in *International Conference on*

Information Communication and Embedded Systems (ICICES2014) (IEEE, 2014), pp. 1–5

20. A.M. Posonia, V.L. Jyothi, XML document retrieval by developing an effective indexing technique, in *2014 Sixth International Conference on Advanced Computing (ICoAC)* (IEEE, 2014), pp. 120–123

21. R.I. Minu, K.K. Thyagharajan, Scrutinizing the video and video retrieval concept. Int. J. Soft Comput. Eng. **1**(5), 270–275 (2011)

22. M. Muaaz, R. Mayrhofer, Smartphone-based gait recognition: from authentication to imitation. IEEE Trans. Mobile Comput. **16**(11), 3209–3221 (2017)

23. M. Shahzad, A.X. Liu, A. Samuel, Secure unlocking of mobile touch screen devices by simple gestures: You can see it but you can not do it, in *Proceedings of 19th Annual International Conference on Mobile Computing Network (MobiCom), New York, NY, USA, 2013*, pp. 3950. https://doi.acm.Org.proxy.library.stonybrook.edu/10.1145/2500423.2500434

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

V. Harsha Vardhan, V. Sri Dileep Kumar, G. Nagarajan, T. Prem Jacob & A. Pravin

Corresponding author

Correspondence to V. Harsha Vardhan .

Editor information

Editors and Affiliations

Sikkim Manipal University, Gangtok, Sikkim, India

Prof. Dr. Karma Sonam Sherpa

Department of Electrical and Electronics Engineering, Sikkim Manipal Institute of Technology, Rangpo, Sikkim, India

Prof. Dr. Akash Kumar Bhoi

Victoria University, Melbourne, VIC, Australia

Prof. Akhtar Kalam

School of Computer Engineering, KIIT University, Bhubaneswar, Odisha, India

Dr. Manoj Kumar Mishra

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Harsha Vardhan, V., Sri Dileep Kumar, V., Nagarajan, G., Prem Jacob, T., Pravin, A. (2021). An Integrated & Robust User Authentication Framework Based on Gait Smart Devices. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_67

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7511-2_67

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5

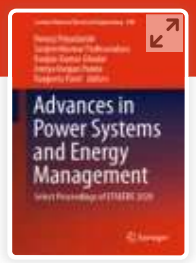
Online ISBN	eBook Packages
978-981-15-7511-2	Energy Energy (R0)

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.



International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy

↳ ETAEERE 2020, ETAEERE 2020: **Advances in Power Systems and Energy Management** pp 531–537

An Extensible Framework for Effective Visualizing and Predicting Crimes Using Big Data Analytics

K. Venkatesh, K. Hari Somasekhar & R. Vignesh

Conference paper | First Online: 21 January 2021

305 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 690)

Abstract

The outcome of the proposed system will help police divisions and law requirement associations to all the more likely comprehend wrongdoing issues and give experiences that will empower them to follow exercises, foresee the probability of occurrences, successfully send assets and improve the dynamic procedure. Crime prediction uses past data and after analyzing data, predict the future crime with location and time. Traditional approaches become less proficient, as their ability to give required outcomes in an opportune way

and inside asset limitations. One promising choice to Crime Investigations is to utilize computational devices dependent on cutting edge information examination to forestall and battle wrongdoing. In this way, machine insight and PC displaying ought to be a necessary piece of the Investigations. Computational, as one of the arrangements, brings quick and effective methods for breaking down information to discover minor proof in enormous and unstructured piles of information.

Keywords

Prediction **Visualization** **Seaborn**

Numpy **Pandas**

This is a preview of subscription content, access via your institution.

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-7504-4_52
- Chapter length: 7 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 245.03

Price includes VAT (India)

- ISBN: 978-981-15-7504-4
- Instant PDF download
- Readable on all devices
- Own it forever

- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

Softcover Book EUR 299.99

Price excludes VAT (India)

- ISBN: 978-981-15-7743-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Softcover Book

Hardcover Book EUR 299.99

Price excludes VAT (India)

- ISBN: 978-981-15-7503-7
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
  Shipping restrictions may apply, check to see if you are impacted.
- Tax calculation will be finalised during checkout

Buy Hardcover Book

Learn about institutional subscriptions

References

1. R. Kumar, A. Pravin, Data protection and outsourcing in cloud with Linear programming and image based OTP. in *2017 International Conference on Information Communication and Embedded Systems (ICICES)* (IEEE, 2017), pp. 1–6

2. B. Schneier, *Tracking the owner of kickass torrents*, https://www.schneier.com/blog/archives/2016/07/trackingtheow.html, July 2017. Accessed July 05 2017

3. D. Griffith, *How to investigate cybercrime*" https://www.policemag.com/channel/technology/articles/2003/11/how-to-investigatecybercrime.aspx, November 2003. Accessed: November 07 2017. O. Guarino, Digital forensics as a big data challenge, in *ISSE 2013 Securing Electronic Business Processes* (Springer, 2013), pp. 197–203

4. T.P. Jacob, K. Pradeep, A multi-objective optimal task scheduling in Cloud environment using Cuckoo Particle Swarm Optimization. Wireless Pers. Commun. **109**(1), 315–331 (2019)

5. L. Graham, *Cybercrime costs the global economy $450 billion: Ceo*, https://www.cnbc.com/2017/02/07/cybercrimecosts-the-global-economy-450-billionceo.html, February 2017. Accessed November 11 2017

6. S. Morgan, *Cyber crime costs projected to reach $2 trillion by 2019*, https://www.forbes.com/sites/stevemorgan/2016/01/17/cyber-crime-costsprojectedto-reach-2-

trillion-by-2019/, January 2017. Accessed November 07 2017

7. Z. Suliman Zubi, A. Altaher Mahmmud, Crime data analysis using data mining techniques to improve crimes prevention. Int. J. Comput. **8**, pp. 39–45 (2014)

8. T. Sonaqwanev, S. Shaikh, S. Shaikh, R. Shinde, A. Sayyad, Crime pattern analysis, visualization and prediction using Data Mining. IJARIIE **1**(4), 681–686 (2015)

9. G. Nagarajan, R.I. Minu, V. Vedanarayanan, S.S. Jebaseelan, K. Vasanth, CIMTEL-Mining Algorithm for Big Data in telecommunication. Int. J. Eng. Technol. **7**(5), 1709–1715 (2015)

10. A. K. Singh, N. Prasad, N. Narkhede, S. Mehta, Crime: classification and pattern prediction. IARJSET **3**(2), 41–43 (2016)

11. K. Chitra Lekha, S. Prakasam, Performance assessment of different classification techniques. CiiT Int. J. Data mining Knowledge Eng. **9**(1), 20–23 (2017)

12. K. Chitra Lekha, S. Prakasam, An analysis of finding the Influencing factors of supporting for the “GiveitUp” LPG subsidy for the

Government using Data mining Techniques. IJCA, **143**(5), 34–39 (2016)

13. A.B. Felicia, L. Lakshmanan, Survey on accident avoidance and privacy preserving navigation system in vehicular network. Global J. Pure. Appl. Math. **12**(1), 943–949 (2016)

14. S. Jancy, C.Jayakumar, Sequence statistical code based data compression algorithm for wireless sensor network. Wireless Personal Commun. **106**(3), 971–985

15. S.L. Jany Shabu, C. Jaya Kumar, Multimodal image fusion and bee colony optimization for brain tumor detection. ARPN J. Eng. Appl. Sci. **13**, 1819–6608 (2018)

16. A. Das, A. Sahi, U. Nandini, SAR image segmentation for land cover change detection, in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, (IEEE, 2016), pp. 1–6

17. P.S. Maran, Meteorological tower wind shear characteristics, vertical wind speed profile and surface roughness analysis near the coastline of Chennai. Songklanakarin J. Sci. Technol. 41(4) (2019)

18. R. Rayan, R. Ramalavanya, A survey on secure data offloading using auction based mechanism, in *2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)* (IEEE, 2017), pp. 335–341

19. A. Ponraj, Optimistic virtual machine placement in cloud data centers using queuing approach. Fut. Gener. Comput. Syst. **93**, 338–344 (2019)

20. G. Lakshmanan, M. Posonia, A novel analysis on application of neural support on nuclear reactor control process monitoring. Indian J. Sci. Technol. **9**, 10 (2016)

21. G. Nagarajan, K.K. Thyagharajan, A machine learning technique for semantic search engine. Procedia Eng. **38**, 2164–2171 (2012)

22. G. Nagarajan, R.I. Minu, Fuzzy Ontology based Multi-Modal semantic information retrieval. Procedia Comput. Sci. **48**, 101–106 (2015)

Author information

Authors and Affiliations

Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

K. Venkatesh, K. Hari Somasekhar & R. Vignesh

Corresponding author

Correspondence to K. Venkatesh .

Editor information

Editors and Affiliations

Birsa Institute Of Technology, Ranchi, Jharkhand, India

Neeraj Priyadarshi

Department of Energy Technology, Aalborg University, Esbjerg, Denmark

Sanjeevikumar Padmanaban

Department of Mechanical Engineering, Sikkim Manipal Institute of Technology, Majitar, Sikkim, India

Ranjan Kumar Ghadai

School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

Amiya Ranjan Panda

School of Electrical Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

Ranjeeta Patel

Rights and permissions

Reprints and Permissions

Copyright information

© 2021 Springer Nature Singapore Pte Ltd.

About this paper

Cite this paper

Venkatesh, K., Somasekhar, K.H., Vignesh, R. (2021). An Extensible Framework for Effective Visualizing and Predicting Crimes Using Big Data Analytics. In: Priyadarshi, N., Padmanaban, S., Ghadai, R.K., Panda, A.R., Patel, R. (eds) Advances in Power Systems and Energy Management. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 690. Springer, Singapore. https://doi.org/10.1007/978-981-15-7504-4_52

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7504-4_52

Published	Publisher Name	Print ISBN
21 January 2021	Springer, Singapore	978-981-15-7503-7

Online ISBN	eBook Packages
978-981-15-7504-4	Energy Energy (R0)

Not logged in - 14.139.184.28

Sathyabama Institute of Science and Technology (2000616985)

SPRINGER NATURE

© 2023 Springer Nature Switzerland AG. Part of Springer Nature.

Smart Innovation, Systems and Technologies 225

Suresh Chandra Satapathy
Vikrant Bhateja
Margarita N. Favorskaya
T. Adilakshmi *Editors*



Smart Computing Techniques and Applications

Proceedings of the Fourth International Conference on Smart Computing and Informatics, Volume 1

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 225

Series Editors

Robert J. Howlett, Bournemouth University and KES International, Shoreham-by-Sea, UK

Lakhmi C. Jain, KES International, Shoreham-by-Sea, UK

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/8767

Suresh Chandra Satapathy · Vikrant Bhateja ·
Margarita N. Favorskaya · T. Adilakshmi
Editors

Smart Computing Techniques and Applications

Proceedings of the Fourth International Conference on Smart Computing and Informatics, Volume 1

Springer

Editors
Suresh Chandra Satapathy
School of Computer Engineering
KIIT University
Bhubaneswar, Odisha, India

Margarita N. Favorskaya
Reshetnev Siberian State University
of Science and Technology
Krasnoyarsk, Russia

Vikrant Bhateja
Department of Electronics
and Communication Engineering
Shri Ramswaroop Memorial Group
of Professional Colleges (SRMGPC)
Lucknow, Uttar Pradesh, India

Dr. A.P.J. Abdul Kalam Technical
University, Lucknow, Uttar Pradesh, India

T. Adilakshmi
Department of Computer Science
and Engineering, Vasavi College
of Engineering, Hyderabad
Telangana, India

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-16-0877-3 ISBN 978-981-16-0878-0 (eBook)
https://doi.org/10.1007/978-981-16-0878-0

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Conference Committees

Chief Patrons

Sri M. Krishna Murthy, Secretary, VAE
Sri P. Balaji, CEO, VCE

Patron

Dr. S. V. Ramana, Principal, VCE

Honorary Chair

Dr. Lakhmi C. Jain, Australia

General Chair

Dr. Suresh Chandra Satapathy, KIIT DU, Bhubaneswar

Organizing Chair

Dr. T. Adilakshmi, Professor and HOD, CSE, VCE

Publication Chairs

Dr. Nagaratna P. Hegde, Professor, CSE, VCE
Dr. Vikrant Bhateja, SRMGPC, Lucknow, UP, India

Program Committee

Dr. S. Ramachandram, Former Vice Chancellor, OU
Dr. Banshidhar Majhi, Director, IITDM Kancheepuram
Dr. Siba K. Udgata, Professor, HCU
Dr. Sourav Mukhopadhyay, Associate Professor, IIT Kharagpur
Dr. P. Radha Krishna, Professor, CSE, NIT Warangal
Dr. M. M. Gore, Professor, MNNIT, Allahabad
Dr. S. M. Hegde, Professor, NIT Surathkal
Dr. S. Bapi Raju, Professor, IIIT Hyderabad
Dr. Rajendra Hegadi, Associate Professor, IIIT Dharwad
Dr. S. Sameen Fatima, Former Principal, OU
Dr. K. Shyamala, Professor, OU
Dr. Naveen Sivadasan, TCS Innovation Labs, Hyderabad
Dr. Badrinath G. Srinivas, Research Scientist—III, Amazon Development Centre, Hyderabad
Dr. Ravindra S. Hegadi, Professor, PAH Solapur University
Dr. S. P. Algur, Professor and Chairman, CSE, Rani Channamma University, Belagavi
Dr. R. Sridevi, HOD, Department of CSE, JNTUH

International Advisory Committee/Technical Program Committee

Dr. Rammohan, South Korea
Dr. Kailash C. Patidar, South Africa
Dr. Naeem Hanoon, Malaysia
Dr. Vimal Kumar, The University of Waikato, New Zealand
Dr. Akshay Sadananda Uppinakudru Pai, University of Copenhagen, Denmark
Dr. K. C. Santosh, The University of South Dakota
Dr. Ayush Goyal, Texas A&M University, Kingsville
Dr. Sobhan Babu, Associate Professor, IIT Hyderabad
Dr. D. V. L. N. Somayajulu, Director, IIIT Kurnool
Dr. Siba Udgata, Professor, HCU
Dr. R. B. V. Subramaanyam, Professor, NITW
Dr. S. G. Sanjeevi, Professor, NITW

Dr. Sanjay Sengupta, CSIR, New Delhi
Dr. A. Govardhan, Rector, JNTU Hyderabad
Prof. Chintan Bhatt, Chandubhai Patel Institute of Technology, Gujarat
Dr. Munesh Chandra Trivedi, ABES Engineering College, Ghaziabad
Dr. Alok Aggarwal, Professor
Dr. Anuja Arora, Jaypee Institute of Information Technology, Noida, India
Dr. Divakar Yadav, Associate Professor, MMMUT, Gorakhpur, India
Dr. Kuda Nageswar Rao, Andhra University, Visakhapatnam
Dr. M. Ramakrishna Murthy, ANITS, Visakhapatnam
Dr. Suberna Kumar, MVGR, Vizianagaram
Dr. J. V. R. Murthy, Director Incubation and IPR, JNTU Kakinada
Dr. D. Ravi, IDRBT, Hyderabad
Dr. Badrinath G. Srinivas, Research Scientist—III, Amazon Development Center, Hyderabad
Dr. K. Shyamala, Professor, OU
Dr. P. V. Sudha, Professor, OU
Dr. M. A. Hameed, Assistant Professor, OU
Dr. B. Sujatha, Assistant Professor, OU
Dr. T. Adilakshmi, Professor and HOD, CSE, VCE
Dr. Nagaratna P. Hegde, Professor, CSE, VCE
Dr. V. Sireesha, Assistant Professor, CSE, VCE

Organizing Committee

Dr. D. Baswaraj, Professor, CSE, VCE
Dr. K. Srinivas, Associate Professor, CSE, VCE
Dr. V. Sireesha, Assistant Professor, CSE, VCE
Mr. S. Vinay Kumar, Assistant Professor, CSE, VCE
Mr. M. Sashi Kumar, Assistant Professor, CSE, VCE
M. Sunitha Reddy, Assistant Professor, CSE, VCE
R. Sateesh Kumar, Assistant Professor, CSE, VCE
Mr. T. Nishitha, Assistant Professor, CSE, VCE

Publicity Committee

Dr. M. Shanmukhi, Professor, CSE, VCE
Mr. C. Gireesh, Assistant Professor, CSE, VCE
Ms. T. Jalaja, Assistant Professor, CSE, VCE
Mr. I. Navakanth, Assistant Professor, CSE, VCE
Ms. S. Komal Kaur, Assistant Professor, CSE, VCE
Mr. T. Saikanth, Assistant Professor, CSE, VCE

Ms. K. Mamatha, Assistant Professor, CSE, VCE
Mr. P. Narasiah, Assistant Professor, CSE, VCE

Website Committee

Mr. S. Vinay Kumar, Assistant Professor, CSE, VCE
Mr. M. S. V. Sashi Kumar, Assistant Professor, CSE, VCE

Preface

This volume contains the selected papers presented at the 4th International Conference on Smart Computing and Informatics (SCI-2020) organized by the Department of Computer Science and Engineering, Vasavi College of Engineering (Autonomous), Ibrahimbagh, Hyderabad, Telangana, during October 9–10, 2020. It provided a great platform for researchers from across the world to report, deliberate, and review the latest progress in the cutting-edge research pertaining to smart computing and its applications to various engineering fields. The response to SCI 2020 was overwhelming with a good number of submissions from different areas relating to artificial intelligence, machine learning, cognitive computing, computational intelligence, and its applications in main tracks. After a rigorous peer review with the help of Technical Program Committee members and external reviewers, only quality papers were accepted for presentation and subsequent publication in this volume of SIST series of Springer.

Several special sessions were floated by eminent professors in cutting-edge technologies such as blockchain, AI, ML, data engineering, computational intelligence, big data analytics and business analytics, and intelligent systems. Eminent researchers and academicians delivered talks addressing the participants in their respective field of proficiency. Our thanks are due to Prof. Roman Senkerik, Head of AI Lab, Tomas Bata University in Zlin, Czech Republic; Shri. Shankarnarayan Bhat, Director Engineering, Intel Technologies India Pvt. Ltd.; Ms. Krupa Rajendran, Associate VP, HCL Technologies; and Mr. Aninda Bose, Springer India, for delivering keynote addresses for the benefit of the participants. We would like to express our appreciation to the members of the Technical Program Committee for their support and cooperation in this publication. We are also thankful to team from Springer for providing a meticulous service for the timely production of this volume.

Our heartfelt thanks to Shri. M. Krishna Murthy, Secretary, VAE; Sri. P. Balaji, CEO, VCE; and Dr. S. V. Ramana, Principal, VCE, for extending support to conduct this conference in Vasavi College of Engineering. Profound thanks to Prof. Lakhmi C. Jain, Australia, for his continuous guidance and support from the beginning of the conference. Without his support, we could never have executed such a mega-event. We are grateful to all the eminent guests, special chairs, track managers, and

reviewers for their excellent support. Special vote of thanks to numerous authors across the country as well as abroad for their valued submissions and to all the delegates for their fruitful discussions that made this conference a great success.

Bhubaneswar, India — Suresh Chandra Satapathy
Lucknow, India — Vikrant Bhateja
Krasnoyarsk, Russia — Margarita N. Favorskaya
Hyderabad, India — T. Adilakshmi

List of Special Sessions Collocated with SCI-2020

SS_01: Next Generation Data Engineering and Communication Technology
Dr. Suresh Limkar, AISSMS Institute of Information Technology, Pune

SS_02: Artificial Intelligence and Machine Learning Applications (AIML)
Dr. V. Sowmya, CEN, Amrita Vishwa Vidyapeetham, Coimbatore
Dr. M. Anand Kumar, NIT Karnataka
Dr. M. Venkatesan, NIT Karnataka
Prof. K. P. Soman, Amrita Vishwa Vidyapeetham, Coimbatore

SS_03: Advances in Computational Intelligence and Its Applications
Dr. C. Kishor Kumar Reddy, Stanley College of Engineering and Technology for Women, Hyderabad
P. R Anisha, Stanley College of Engineering and Technology for Women, Hyderabad

SS_04: Blockchain Technology: Foundations, Challenges and Applications
Prof. Sandeep Kumar Panda, Faculty of Science and Technology, ICFAI Foundation for Higher Education, Hyderabad
Prof. Santosh Kumar Swain, School of Computer Engineering, KIIT (Deemed to be) University, Bhubaneswar

SS_05: Application of Machine Learning for Intelligent System Design
Dr. Minakhi Rout, KIIT (Deemed to be) University, Bhubaneswar

SS_06: Advances in Big Data Analytics and Business Intelligence
Dr. Vijay B. Gadicha, G. H. Raisoni Academy of Engineering and Technology Nagpur
Dr. Ajay B. Gadicha, P. R. Pote College of Engineering and Management, Amravati

SS_07: Recent Advances in Artificial Intelligence—Applications, Challenges and Future Trends

Dr. S. Velliangiri, CMR Institute of Technology, Hyderabad
Dr. P. Karthikeyan, Presidency University, Bengaluru
Dr. Iwin Thanakumar Joseph, KITS, Coimbatore

Contents

1 Pad Vending Machine with Cashless Payment 1
Anjali Prajapati and Anandu M. Dharan
1.1 Introduction 1
1.2 Literature Review 2
1.3 Proposed System 3
1.4 Implementation 4
1.5 Results and Comparative Study 5
1.6 Future Scope 6
1.7 Conclusion 6
References 7

2 Design of Wrung Order Layout Using Advanced Optimization Techniques with Integrated Variable Batch Scheduling 9
K. Mallikarjuna and Y. Hariprasad Reddy
2.1 Introduction 9
2.2 Problematic Depiction 10
2.2.1 Multi-goal Mathematical Representations 10
2.3 Proposed Advanced Optimization 11
2.3.1 Sheep Run Technique 11
2.3.2 Tabu Search 12
2.4 Results and Discussion 13
2.4.1 Inferences 14
2.5 Conclusion 14
References 16

3 Adaptive Beamforming Using Radial Basis Function 17
Gowtham Narayan, Sharath Kumar, Abhinav Bharadwaj, Srikar Vangala, and R. Lavanya
3.1 Introduction 17
3.2 Methodology 18
3.3 RBFNN Overview 20

3.4 Results ... 22
3.5 Conclusion ... 22
References ... 24

4 A Cloud-Based Centralized Smart Electricity Billing and Payment Management System ... 25
P. Pramod Kumar, Nethaji Achha, G. Ranadheer Reddy, and V. Pranathi
4.1 Introduction ... 25
4.2 Related Work ... 26
4.3 Methodology ... 28
4.3.1 Automatic or Calculation Section ... 28
4.3.2 Payment Section ... 29
4.3.3 Rural Areas Billing ... 30
4.4 Performance Analysis ... 30
4.4.1 Comparison of Smart Meter with Traditional Meters ... 30
4.4.2 Research Limitations and Practical Implications ... 31
4.4.3 Originality and Outcomes ... 31
4.5 Conclusion and Future Scope ... 32
References ... 32

5 Real-Time Driver Distraction Detection Using OpenCV and Machine Learning Algorithms ... 35
V. N. V. L. S. Swathi, D. Akhilesh, G. Senthil Kumar, and A. Vani Vathsala
5.1 Introduction ... 35
5.2 Background ... 36
5.2.1 Real-Time System [1] ... 36
5.2.2 Driver Distraction Detection System [2] ... 36
5.2.3 OpenCV ... 36
5.3 Related Study ... 37
5.4 Methodology ... 37
5.4.1 Distraction Prevention Before Distraction Occurs ... 38
5.4.2 Distraction Mitigation After Distraction Occurs ... 43
5.4.3 Collision System Adjustment When a Potential Collision ... 43
5.5 Implementation ... 43
5.6 Result ... 44
5.7 Conclusion and Future Work ... 46
References ... 46

6 Data-Driven Prediction Model for Crime Patterns ... 47
Y. Gayathri, Y. Sri Lalitha, M. V. Aditya Nag, and Sk. Althaf Hussain Basha
6.1 Introduction ... 47
6.2 Literature Survey ... 48

6.3 Our Approach 49
6.4 Experimental Approach 50
6.4.1 Result Analysis by Visualization 51
6.4.2 Construction of Prediction Models 54
6.5 Evaluation 57
6.6 Conclusions 57
References 58

7 Automatic Reminder List Generation to Assist People with Alzheimer's 59
Shireesha Gundeti and Bhageshwari Ratkal
7.1 Introduction 59
7.2 Existing Systems 60
7.3 Proposed System—Automatic Generation of Reminder List (AGRL) 61
7.3.1 User Input 62
7.3.2 AGRL Processing 62
7.3.3 The User Gets Notified of All the Events Recorded as Audio or Text 62
7.4 AGRL Processing 62
7.4.1 Text—Preprocessing 62
7.4.2 Record—Preprocessing 64
7.4.3 Text and Record—Processing 64
7.5 Discussions on Results 66
7.6 Conclusion 70
References 70

8 Recognizing the Faces from Surveillance Video 73
T. Shreekumar, K. Karunakara, H. Manjunath, B. Ravinarayana, and D. R. Annappa Swami
8.1 Introduction 74
8.2 Literature Survey 75
8.3 Face Recognition with LLR and ANN 77
8.4 Face Recognition with PSO and SVM 78
8.5 Experiment and Result 79
8.6 Conclusion 84
References 85

9 Application of Bat Optimization Algorithm for Power System Loss Reduction Using DSSSC 87
M. Thirumalai and T. Yuvaraj
9.1 Introduction 87
9.2 Formulation of the Problem 88
9.2.1 Analysis of Load Flow 88
9.2.2 Power Loss Mitigation 90

9.3 Proposed Bat Algorithm ... 90
9.3.1 Overview of Bat Algorithm ... 90
9.3.2 Population Initialization ... 90
9.3.3 Motion of Virtual Bats ... 91
9.3.4 Emission of Pulse and Loudness ... 91
9.4 Discussion on Results ... 92
9.4.1 Test System for IEEE 69-Bus ... 92
9.5 Conclusion ... 94
References ... 94

10 Malaria Cell Detection Model ... 95
Prasannajeet Bajpai and Sanchita Chourawar
10.1 Introduction ... 95
10.2 Literature Review ... 96
10.3 Architecture and Working ... 97
10.4 Result Analysis ... 101
10.5 Conclusion ... 101
References ... 101

11 Design and Analysis of Fractal Monopole Antennas for Multiband Wireless Applications ... 103
Siddhi Oja and P. M. Menghal
11.1 Introduction ... 103
11.2 Fractal Monopole Antennas ... 104
11.3 Antena Design ... 104
11.4 Simulation Results and Comparative Analysis ... 105
11.5 Conclusion ... 109
Appendix 'A' ... 115
References ... 116

12 Fake News Detection Using Text Analytics ... 117
Uma Maheshwar Amanchi, Nithesh Badam, and Rama Lakshmi Elaganti
12.1 Introduction ... 117
12.2 Proposed System ... 118
12.2.1 Module 1 ... 118
12.2.2 Module 2 ... 118
12.3 Implementation ... 119
12.3.1 FND (Fake News Detection System) ... 120
12.3.2 Data Preprocessing ... 121
12.3.3 TF-IDF Vectorization ... 121
12.3.4 Clustering ... 121
12.3.5 Topic Modeling ... 122
12.3.6 Evaluation Credibility ... 122
12.3.7 Performance Evaluation ... 122

12.4 Results ... 123
12.4.1 Cluster Evaluation ... 123
12.4.2 Accuracy of FND System ... 123
12.5 Conclusion ... 124
References ... 125

13 An Extreme Learning Machine-Based Model for Cryptocurrencies Prediction ... 127
Sarat Chandra Nayak, B. Satyanarayana, Bimal Prasad Kar, and J. Karthik
13.1 Introduction ... 127
13.2 Methods and Materials ... 128
13.2.1 ELM ... 129
13.2.2 Proposed ELM-Based Prediction ... 130
13.2.3 Cryptocurrency Data ... 131
13.3 Experimental Results and Analysis ... 131
13.4 Conclusions ... 132
References ... 135

14 An Ensemble Approach for Intrusion Detection in Collaborative Attack Environment ... 137
Shraddha R. Khonde, G. Kulanthaivel, and V. Ulagamuthalvi
14.1 Introduction ... 137
14.2 Literature Survey ... 138
14.3 Methodology of Proposed System ... 140
14.3.1 Signature Creation for Collaborative Attack Over UNSW NB15 ... 140
14.3.2 Feature Selection ... 141
14.3.3 Proposed System Architecture ... 141
14.4 Results and Discussions ... 142
14.5 Conclusion ... 144
References ... 145

15 A Decision Based Asymmetrically Trimmed Modified Geometric Mean Algorithm for the Removal of High Density Salt and Pepper Noise in Images and Videos ... 147
K. Vasanth, C. N. Ravi, S. Nagaraj, M. Vadivel, K. Gopi, S. Pradeep Kumar Reddy, and Thulasi Prasad
15.1 Introduction ... 148
15.2 Decision-Based Neighborhood Referred Unsymmetrical Trimmed Modified Geometric Mean (DBNRUTMGMA) ... 148
15.3 Simulation Results ... 149
15.4 Conclusion ... 153
References ... 154

16 Anti-poaching System Using Wireless Sensor Networking 155
T. Asha, H. S. Jahanavi, Sony S. Rao, S. Sinchana, and K. A. Priyanka
16.1 Introduction 155
16.2 Literature Survey 156
16.3 System Design 157
16.4 System Implementation 159
16.4.1 Flame Sensor 159
16.4.2 Vibration Sensor 160
16.4.3 Metal Detector Sensor 160
16.4.4 Radio-Frequency Identification (RFID) 160
16.4.5 Bylnk App 161
16.4.6 Arduino Ide 161
16.5 Results 161
16.6 Conclusion 161
References 163

17 Performance Issues and Monitoring Mechanisms for Serverless IoT Applications—An Exploratory Study 165
Shajulin Benedict
17.1 Introduction 165
17.2 Related Works 166
17.3 Serverless IoT-Working Model 167
17.4 Performance Monitoring Mechanisms 168
17.5 Monitoring Tools—A Comparative Study 170
17.5.1 Challenges of Tools—Research Directions 171
17.6 Conclusion 173
References 174

18 Detection of Pemphigus Vulgaris in Development Stage of Skin Erosion 175
Mayuri Vagh and Dipali Kasat
18.1 Introduction 175
18.2 Literature Survey 176
18.3 System Overview 177
18.3.1 Primary Section/ Interrogate About Symptoms 177
18.3.2 Secondary Section/ Image Analysis 178
18.4 Result Analysis 180
18.5 Conclusion 182
References 182

19 Anomaly-Based Intrusion Detection System in Two Benchmark Datasets Using Various Learning Algorithms 185
Thongam Jayaluxmi Devi and Khundrakpam Johnson Singh
19.1 Introduction 185
19.2 Related Works 186

19.3 Datasets ... 186
19.3.1 KDD Cup 99 ... 186
19.3.2 NSL-KDD ... 186
19.4 Proposed Methodology ... 188
19.4.1 The Classifiers ... 191
19.5 Results and Discussion ... 191
19.6 Conclusion ... 193
References ... 193

20 A Machine Learning Approach to Intrusion Detection System Using UNSW-NB-15 and CICDDoS2019 Datasets ... 195
Akoijam Priya Devi and Khundrakpam Johnson Singh
20.1 Introduction ... 195
20.2 Related Works ... 196
20.3 Dataset ... 196
20.3.1 UNSW-NB-15 Dataset ... 197
20.3.2 CICDDoS2019 Dataset ... 197
20.4 Proposed Methodology ... 198
20.4.1 J48 Classifier ... 199
20.4.2 Naïve Bayes Classifier ... 200
20.4.3 Decision Table Classifier ... 200
20.4.4 Simple Logistic Based Classifier ... 200
20.5 Results and Discussions ... 200
20.6 Conclusions and Future Scope ... 202
References ... 204

21 Discovery of Localized Malicious Attack in Wireless Networks ... 207
S. Gowri, Senduru Srinivasulu, J. Jabez, J. S. Vimali, and A. Sivasangari
21.1 Introduction ... 207
21.2 Literature Survey ... 208
21.3 Proposed Methodology ... 209
21.3.1 Detection Model of Generalized Attack ... 210
21.4 Results and Discussion ... 212
21.5 Screenshots ... 212
21.6 Conclusion and Future Work ... 212
References ... 216

22 An Adapted Approach of Image Steganography Using Pixel Mutation and Bit Augmentation ... 217
Ravi Saini, Kamaldeep Joshi, and Rainu Nandal
22.1 Introduction ... 217
22.2 Proposed Method ... 218
22.3 Algorithm ... 219
22.3.1 Insertion Algorithm ... 219
22.3.2 Retrieval Algorithm ... 220
22.4 Results and Analysis ... 220

22.5 Conclusion and Future Scope 222
References 224

23 CNN Architectures to Recognize Handwritten Telugu Characters 225
K. Dhana Sree Devi and C. Shoba Bindu
23.1 Introduction 225
23.2 Architectural Study of CNN 226
23.2.1 General Artificial Neural Network 226
23.2.2 Convolution Neural Networks 229
23.3 Handwritten Telugu Character Recognition 235
23.3.1 Classical LeNet-5 Architecture 235
23.3.2 Proposed ConvNet Architecture 235
23.3.3 Data Preparation 236
23.4 Experimental Results 237
23.5 Conclusions 237
References 238

24 Rider-Deep Belief Network-Based MapReduce Framework for Big Data Classification 241
Sridhar Gujjeti and Suresh Pabboju
24.1 Introduction 241
24.2 Literature Review 242
24.3 Proposed Rider-DBN for Predicting Disease Using Big Data 242
24.3.1 Selection of Imperative Feature Using Information Gain in Mapper Phase 243
24.3.2 Classification of Massive Data with Proposed Rider-DBN in Reducer Phase 244
24.4 Results and Discussion 246
24.4.1 Comparative Analysis 247
24.5 Conclusion 249
References 249

25 Modified Compressive Sensing and Differential Operation Based Channel Feedback Scheme for Massive MIMO Systems for 5G Applications 251
V. Baranidharan, R. Shalini, N. Nithya Shree, S. Sasithra, and R. Jaya Murugan
25.1 Introduction 252
25.2 Literature Survey 253
25.3 System Model 253
25.3.1 Temporal Correlation of Time-Varying MIMO Channels 254
25.3.2 Proposed Modified Differential Channel Feedback Based Compressive Sensing 255

25.4 Simulated Results 256
25.4.1 Channel Impulse Response Comparison 256
25.4.2 NMSE Validation 258
25.5 Conclusion 258
References 260

26 Smart Shopping Bag Using IoT 261
M. Suguna, R. V. Kanimozhi, D. Naveen Kumar, S. Indhushree, and D. Prakash
26.1 Introduction 262
26.2 Literature Review 262
26.3 Purpose and Objective 264
26.4 Proposed System 264
26.5 Module Description 265
26.5.1 RFID Reader Module 265
26.5.2 LCD Display Module 265
26.5.3 Switch Module 266
26.5.4 Blind People Module 266
26.5.5 Band Module 266
26.5.6 Billing Unit 266
26.6 Results and Discussion 267
26.7 Conclusion 269
References 269

27 An Efficient VM Live Migration Through VM I/O Requests Using Live Scheduling Scheme (LSS) 271
M. N. Faruk, K. Lakshmi Prasad, and G. Lakshmi Vara Prasad
27.1 Introduction 272
27.1.1 Definitions 273
27.2 Related Works 274
27.3 Proposed Method 275
27.3.1 I/O Threads Level Interference 276
27.3.2 System Execution and Workflow 277
27.4 Experimental Analysis and Discussions 281
27.5 Conclusion 283
References 284

28 Real-Time Area-Based Traffic Density Calculation Using Image Processing for Smart Traffic System 287
S. Rakesh and Nagaratna P. Hegde
28.1 Introduction 287
28.2 System Model 289
28.3 Implementation 289
28.4 Results 291
28.5 Conclusion and Future Scope 292
References 295

29 An Association Mining Rules Implemented in Data Mining 297
B. Varija and Nagaratna P. Hegde
29.1 Introduction 297
29.2 The Association Rules 298
29.3 Apriori Algorithm 298
29.4 Multi-level Associative Rule 299
29.5 Multidimensional Association Rule 300
29.5.1 Mining Using Dynamic Discretization 300
29.5.2 Distance-Based Association Rule 300
29.6 Fuzzy Association Rule 301
29.7 Hybrid Dimensional Association Rules 301
29.8 Weighted Association Rule Mining 301
29.9 Association Rule Mining in Different Databases 301
29.10 Conclusion 302
References 303

30 Wi-Fi Networking: A Novel Method for Stabilizing Internet Connectivity by Optimizing Wi-Fi Router Configuration 305
Vijay A. Kanade
30.1 Introduction 305
30.2 Wi-Fi Networking 306
30.3 Radio Waves 307
30.3.1 RF System Components 308
30.4 Research Developments 310
30.5 Optimizing Wi-Fi Router Configuration 311
30.5.1 Use Case 312
30.5.2 Algorithm 313
30.6 Advantages 313
30.7 Conclusion 313
References 314

31 Periocular Segmentation Using K-Means Clustering Algorithm and Masking 315
V. Sandhya and Nagaratna P. Hegde
31.1 Introduction 315
31.2 Related Work 316
31.3 Median Filtering 317
31.4 Canny Edge Detection Algorithm 317
31.5 K-Means Clustering Algorithm 317
31.6 Masking 317
31.7 Proposed Method 318
31.8 Data Sets 319
31.8.1 UBIRIS.V1 319
31.8.2 UBIRIS.V2 319
31.8.3 UBIPr 319
31.8.4 Visob 320

31.9 Results ... 320
31.10 Conclusion ... 320
References ... 321

32 RPCR: Retransmission Power-Cost Aware Routing in Mobile Ad Hoc Networks ... 323
V. Sowmya Devi and Nagaratna P. Hegde
32.1 Introduction ... 323
32.2 Literature Review ... 324
32.3 Power-Cost Aware Routing ... 325
32.4 Retransmission-Power-Cost Aware Routing (RPCR) ... 326
32.5 Simulation Results ... 328
32.6 Conclusion ... 329
References ... 330

33 Comparison on Different Filters for Performance Improvement on Fingerprint Image Enhancement ... 333
Meghna B. Patel, Jagruti N. Patel, Satyen M. Parikh, and Ashok R. Patel
33.1 Introduction ... 333
33.2 Related Work ... 334
33.3 Proposed Research Work ... 336
33.3.1 Mean Filter ... 336
33.3.2 Min Filter ... 336
33.3.3 Max Filter ... 336
33.3.4 Median Filter ... 337
33.3.5 Sobel Edge Detection [10, 22] ... 337
33.3.6 Gabor Filter [13, 24, 25] ... 338
33.3.7 O'Gorman Filter [13] ... 338
33.4 Evaluation Parameter and Result Discussion ... 339
33.4.1 Quality of an Image ... 339
33.4.2 PSNR and MSE Value ... 339
33.4.3 Computational Time ... 341
33.5 Conclusion ... 342
References ... 342

34 Framework for Spam Detection Using Multi-objective Optimization Algorithm ... 345
M. Deepika and Nagaratna P. Hegde
34.1 Introduction ... 345
34.2 Background Work ... 347
34.3 Proposed Model ... 348
34.4 Results and Discussion ... 352
34.5 Conclusion ... 354
References ... 354

35 Time Series Analysis of Assam Rainfall Using SARIMA and ARIMA 357
Utpal Barman, Asif Ekbal Hussain, Mridul Jyoti Dahal, Puja Barman, and Mehnaz Hazarika
35.1 Introduction 357
35.2 Material and Methods 358
35.2.1 About Dataset 358
35.2.2 Preprocessing and Components of Rainfall Data 358
35.2.3 Time Series Analysis of Rainfall Data Using ARIMA and SARIMA 360
35.3 Results and Discussion 362
35.4 Conclusion 363
References 363

36 Reliable Low-Cost High-Speed and Direction-Controlled Surveillance Robot 365
S. Rooban, I. Venkata Sai Eshwar, R. Radhika, M. Dheeraj Kumar, and K. Chaitanya Krishna
36.1 Introduction 366
36.2 Related Work 366
36.3 Proposed Methodology 367
36.3.1 Patrolling Unit 367
36.3.2 Vision Unit 369
36.3.3 Environmental Identification Unit 371
36.4 Results 371
36.5 Conclusion 373
References 374

37 Deciphering WEP, WPA, and WPA2 Pre-shared Keys Using Fluxion 377
Sidharth Atluri and Revanth Rallabandi
37.1 Introduction 377
37.1.1 WEP 377
37.1.2 WPA/WPA2 378
37.2 Related Work 378
37.3 Existing System 378
37.4 Proposed System 380
37.4.1 Installation 380
37.4.2 Scanning for Wi-Fi Networks 381
37.4.3 Selecting Target Access Point 382
37.5 Results 384
37.6 Conclusion 384
References 385

38 Blockchain-Based Government Project Fund Tracking and Management System 387
M. V. Varada, Wissam Salih Abdulla, Aparna Udayakumar, Haritha Raghu, and Manoj V. Thomas
38.1 Introduction 387
38.2 Related Work 388
38.2.1 Application of Blockchain in Government 388
38.2.2 Hyperledger Fabric 388
38.3 Proposed System 389
38.3.1 Architecture of the System 389
38.3.2 Specialized Network 390
38.3.3 Storage 390
38.3.4 Requirements 391
38.3.5 Working of the System 391
38.4 Modular Description 391
38.5 Results and Discussion 393
38.6 Conclusion and Future Work 394
References 394

39 Random Valued Impulse Noise Detection Using Fuzzy c-Means Clustering Technique 397
Aritra Bandyopadhyay, Kaustuv Deb, Atanu Das, and Rajib Bag
39.1 Introduction 397
39.2 Proposed Methodology 399
39.3 Results and Discussion 400
39.4 Conclusion 404
References 404

40 Distress-Level Detection Using Deep Learning and Transfer Learning Methods 407
Srijha Kalyan, Hamsini Ravishankar, and C. Arunkumar
40.1 Introduction 407
40.2 Related Methods 408
40.3 Proposed Approach 409
40.3.1 Dataset and Feature Selection 409
40.3.2 Methodology 410
40.3.3 Transfer Learning Models Implemented 411
40.3.4 Ensemble Methodology in Flask Application 411
40.4 Results and Discussion 411
40.4.1 Model Performance Analysis 411
40.5 Conclusion and Future Work 413
References 414

41 Fingerprint Genuinity Classification 415
Meega Annu Jacob and Nirmal Varghese Babu
41.1 Introduction 415
41.2 Related Works 416
41.3 Proposed System 416
41.4 Experimental Results 419
41.5 Conclusion and Future Work 420
41.6 Declaration 421
References 421

42 Feature Selection by Associativity for Sentiment Analysis 423
S. Fouzia Sayeedunnisa, Nagaratna P. Hegde, and Khaleel Ur Rahman Khan
42.1 Introduction 423
42.2 Related Work 424
42.3 Feature Selection 425
42.3.1 Existing Feature Selection Methods 425
42.3.2 Proposed Method Using Feature Association as a Feature 426
42.3.3 Preprocessing 426
42.3.4 Sentiment Lexicon Extraction 426
42.4 Classifier 428
42.5 Experimental Study 428
42.5.1 Datasets and Statistics 428
42.6 Conclusion 429
References 430

43 Optimization of Serviceability of Mobile Communication Towers Under Capital Constraints 431
Pankaj Preet Singh Dhaliwal, Mahendra Kumar, and Veerendra Dakulagi
43.1 Introduction 431
43.2 Points to Be Observed for Maintaining High Degree of Serviceability 432
43.3 Spare Parts Management 433
43.4 Assurance Levels 434
43.5 Quantity of Modules/Components Required to Be Stocked for Desired Assurance Levels 434
43.6 Conclusion 437
References 438

44 KruthVidya—Mobile App to Detect Dyslexia 441
Apurva Dasari, Rashmitha Proddaturi, Sai Prasanna Vasaraju, and Kalyani Nara
44.1 Introduction 441
44.2 Epidemics of Dyslexia 442

44.3 Factors in Phonological Theory and Their Relation to Dyslexia 443
44.3.1 Access Versus Representations 443
44.3.2 Causal Direction 443
44.3.3 Orthographic Learning 444
44.4 Motivation 444
44.5 Architecture of KruthVidya 446
44.5.1 Collection of Inputs 448
44.5.2 Implementation of OCR to Extract Text 448
44.5.3 Speech-To-Text Conversion of the Audio Input Collected 449
44.5.4 Output Results 451
44.6 Conclusion 452
References 452

45 A Survey on Streaming Adaptation Techniques for QoS and QoE in Real-Time Video Streaming 455
Satyanarayana Reddy Marri and P. Chenna Reddy
45.1 Introduction 456
45.2 Motivation 456
45.3 Overview of Adaptive Streaming Techniques and Video Coding Standards 457
45.4 QoS and QoE in Real-Time Data Streaming 459
45.5 Machine Learning Techniques 460
45.6 State of the Art on Video Streaming Adaptation Algorithms in QoS and QoE 461
45.7 Conclusion 461
References 463

46 A Study on the State of Artificial Intelligence in the Design Phase of Software Engineering 467
Mriganka Shekhar Sarmah, Jediael Meshua Sumer, Marlom Bey, and Bobby Sharma
46.1 Introduction 467
46.2 Related Works 468
46.3 Results and Analysis of Existing Work 471
46.4 Conclusion and Future Work 472
References 472

47 S2S Translator by Natural Language Processing 475
Srikanth Bethu, Suresh Mamidisetti, S. Bhargavi Latha, and B. Sankara Babu
47.1 Introduction 475
47.2 Literature Review 476
47.3 Methodology and Implementation 477
47.4 Results and Discussion 481

47.5 Conclusion 481
References 484

48 Chemical Plant Liquid Leakage IoT-Based Monitoring 487
G. P. Hegde, Nagaratna Hegde, and M. Seetha
48.1 Introduction 487
48.1.1 Liquid Superiority Monitoring 487
48.2 Earlier Work 489
48.2.1 Real-Time Work 489
48.2.2 Liquid Quality Monitoring 489
48.3 Existing Framework 490
48.3.1 Real-Time Work 490
48.4 Proposed Framework 490
48.4.1 System Overview 490
48.4.2 System Components 492
48.4.3 Internal Overview 493
48.5 Results and Discussions 494
48.5.1 Implementation 494
48.6 Conclusion 495
References 495

49 K-Hop Iterative Time Synchronization Over a Linear Connected Wireless Sensor Networks 497
K. L. V. Sai Prakash Sakuru and N. Bheema Rao
49.1 Introduction 497
49.2 Iterative Clock Synchronization Algorithm 499
49.3 Two-Hop Synchronization 500
49.4 K-Hop Synchronization 501
49.5 Proposed Algorithm—Modified Two-Hop Synchronization 502
49.6 Proposed Algorithm—Modified K-Hop Synchronization 503
49.7 Simulations 504
49.8 Conclusions and Future Work 505
References 505

50 A Convolutional Neural Network Architecture for Tomato Leaf Disease Detection Using Data Augmentation 507
Matta Bharati Devi and K. Amarendra
50.1 Introduction 507
50.2 Related Work 508
50.3 Methodology 509
50.3.1 Data Preprocessing 510
50.3.2 Data Augmentation 510
50.3.3 Convolutional Neural Network for Tomato Leaf Disease Identification 511
50.3.4 Training Details 512

50.4 Experimental Results 512
50.4.1 Tomato Leaf Disease Dataset 512
50.4.2 Performance Evaluation Measures 512
50.4.3 Results Analysis 513
50.5 Conclusion 514
References 515

51 Development of Robust Framework for Automatic Segmentation of Brain MRI Images 517
K. Bhima and A. Jagan
51.1 Introduction 517
51.2 Development of Robust Framework 518
51.3 Results and Discussion 521
51.4 Conclusion 523
References 523

52 Yield Estimation and Drought Monitoring Through Image Processing Using MATLAB 525
Shaikh Akbar Shaikh Rasul, Jadhav Swamini Narendra, and Dipti Y. Sakhare
52.1 Introduction 525
52.2 Proposed Algorithm 526
52.2.1 Main Building Blocks 526
52.2.2 Data Acquisition 526
52.2.3 Preprocessing of Data 527
52.2.4 Data Processing and MATLAB Programming 527
52.2.5 Explanation of MATLAB Code 528
52.3 Experiment and Result 530
52.3.1 Results of Experiment Through MATLAB 530
52.3.2 Tables and Quantitative Analysis of Results 532
52.3.3 Graphical Analysis of Result 533
52.3.4 Result of Another Experiment on 10 Images 533
52.4 Conclusion 535
References 536

53 A Framework for Efficient Multilevel Polarity-Based Sentiment Analysis Using Fuzzy Logic 537
Shiramshetty Gouthami and Nagaratna P. Hegde
53.1 Introduction 537
53.2 Classification of Sentiment Analysis 539
53.2.1 Document Level of Sentiment Analysis 539
53.2.2 Sentence Level of Sentiment Analysis 539
53.2.3 Aspect-Based Sentiment Analysis 540
53.2.4 Comparative Sentiment Analysis 540
53.2.5 Sentiment Lexicon Acquisition 540

53.3 Mutilevel Polarity Sentiment Analysis Using Fuzzy Logic 541
53.3.1 Online Reviews 542
53.3.2 Sentence Separation Using SentiWordNet 542
53.3.3 Sentence Label Propagation 543
53.3.4 Sentiment Classifier 543
53.3.5 Fuzzy Logic Interface System 544
53.3.6 Fuzzy Sets 545
53.4 Results 545
53.5 Conclusion 547

54 Practical Assessment of Application Layer Protocols for Resource Constraint IoT Applications 549
Manasi Mishra and S. R. N. Reddy
54.1 Introduction 549
54.2 IoT Application Layer Protocols 551
54.2.1 HTTP 551
54.2.2 MQTT 552
54.2.3 CoAP 552
54.2.4 AMQP 553
54.2.5 WebSocket 553
54.3 Experimental Test Bed Setup 556
54.3.1 Scenario 556
54.3.2 Key Performance Indicator 556
54.3.3 Experiment Flow Chart 557
54.4 Results and Discussion 558
54.5 Conclusion and Future Work 558
References 559

55 Collaborative-Based Movie Recommender System—A Proposed Model 561
Prajna Paramita Parida, Mahendra Kumar Gourisaria, Manjusha Pandey, and Siddharth Swarup Rautaray
55.1 Introduction 561
55.2 Literature Survey 563
55.3 Proposed Model 565
55.4 Feature Selection 567
55.5 Implementation 567
55.6 Simulation Results 568
55.7 Conclusion and Future Work 570
References 570

56 Classification of Bruised Apple Using Ultrasound Technology and SVM Classifier 573
Gopinath Bej, Tamal Dey, Abhra Pal, Sabyasachi Majumdar, Rishin Banerjee, Devdulal Ghosh, Vamshi Krishna Palakurthi, Amitava Akuli, and Nabarun Bhattacharyya

56.1 Introduction 574
56.2 Methodology 575
56.2.1 Sample Collection 575
56.2.2 Ultrasonic Setup 575
56.2.3 Signal Acquisition 576
56.2.4 Manual Analysis 577
56.3 Data Analysis 577
56.4 Results and Discussion 578
56.5 Conclusion and Future Work 579
References 581

57 An Artificial Neural Networks Feature Extraction Approach to Predict Nephrolithiasis (Kidney Stones) Based on KUB Ultrasound Imaging 583
Gollapalli Sumana, Giri Aparna, and Gade Anitha Mary
57.1 Introduction 583
57.1.1 Methodology 584
57.2 Automatic Extraction of KUB Ultrasound Image Features Using Digital Image Processing Techniques 585
57.2.1 Introduction 585
57.2.2 Data 586
57.2.3 Procedure 586
57.2.4 Extraction of Ultrasound Image Features 586
57.3 Predictions of Nephrolithiasis Based on Ultrasound Extracted Features Using Neural Networks 590
57.3.1 Introduction 590
57.3.2 Parameters 591
57.3.3 Procedure 591
57.4 Analysis and Results 591
57.5 Conclusion 593
References 596

58 Multi-resource Task Scheduling for Minimum Task Completion Time in Cloud Computing Using Credit-Based Assignment Problem 597
Animoni Nagaraju and Y. Rama Devi
58.1 Introduction 597
58.1.1 Task Scheduling in Cloud Computing 599
58.2 Proposed Methodology 600
58.3 Algorithm of Multi-Resource Credit-Based Assignment Problem with n Tasks 600
58.4 Lexi Search Procedure of Multiple Resources Credit-Based Two Appliance n Task Assignment Problem 601
58.5 Computational Experience 605
58.6 Conclusion 606
References 606

59 Insights into the Advancements of Artificial Intelligence and Machine Learning, the Present State of Art, and Future Prospects: Seven Decades of Digital Revolution 609
Prabhleen Bindra, Meghana Kshirsagar, Conor Ryan, Gauri Vaidya, Krishn Kumar Gupt, and Vivek Kshirsagar
59.1 Introduction 610
59.2 The Evolving Hype of AI/ML 610
59.2.1 The Past 610
59.2.2 The Present 611
59.3 Subdomains of AI/ML 611
59.3.1 Artificial Neural Networks (ANN) 611
59.3.2 Evolutionary Computation 612
59.3.3 Natural Language Processing (NLP) 612
59.3.4 Bayesian Networks and Fuzzy Logic 612
59.3.5 Internet of Things (IoT) 613
59.4 Popular AI Methods 613
59.4.1 Supervised Learning 613
59.4.2 Unsupervised Learning 615
59.5 Market Trends with AI Applications 615
59.5.1 AI Use Cases 615
59.5.2 Popular AI-Enabled Products 617
59.6 Future with AI 617
59.6.1 Emerging Trends for the Future 617
59.7 Conclusions 618
References 619

60 Development of Porting Analyzer to Search Cross-Language Code Clones Using Levenshtein Distance 623
Sanjay B. Ankali and Latha Parthiban
60.1 Introduction 624
60.2 Background and Related Work 624
60.2.1 Background 625
60.2.2 Related Work 626
60.3 Proposed Methodology 627
60.4 Results and Discussion 628
60.5 Conclusion 630
References 631

61 Multi-Objective Genetic Algorithm for Hyperspectral Image Analysis 633
T. Subba Reddy, J. Harikiran, and B. Sai Chandana
61.1 Introduction 633
61.2 GA Based Hyperspectral Image Contrast Enhancement 634
61.3 Dimensionality Reduction of Hyperspectral Dataset Using Genetic Algorithm 636
61.4 GA Based Hyperspectral Image Classification 638

61.5 Experimental Results 638
61.6 Conclusions 641
References 641

62 Auto-Discovery and Monitoring of Network Resources: SNMP-Based Network Mapping and Fault Management 643
Shaik Imam Saheb and Abdul Rasool Md
62.1 Introduction 643
62.1.1 Network Management Architecture 644
62.2 Related Work 645
62.2.1 Internet Control Message Protocol (ICMP) 645
62.3 SNMP Architecture 646
62.3.1 SNMP Agents and Managers 646
62.3.2 SMNP Data Structures 646
62.4 Design and Implementation 647
62.4.1 Auto-Discovery and Network Monitoring Using ICMP and SNMP 647
62.5 Discussion 649
62.6 Conclusion 649
References 653

63 Privacy Preserving Semantic-Aware Multi-keyword Ranked Dynamic Search Over Encrypted Data 655
M. D. Asrar Ahmed, S. Ramachandram, and Khaleel Ur Rahman Khan
63.1 Introduction 655
63.2 Related Work 656
63.3 Proposed Scheme: Notations, Semantic Search, and Algorithms 657
63.3.1 Notations 657
63.3.2 Semantic Extension of Queries—Word2Vec Model 658
63.3.3 Proposed Scheme 658
63.4 Security and Performance Analysis of Proposed Scheme 661
63.4.1 Security Analysis 661
63.4.2 Performance Analysis 661
63.5 Conclusion 662
References 664

64 A Thorough Study on Weblog Files and Its Analysis Tools 665
N. Vanitha and M. Suriakala
64.1 Introduction 665
64.2 Web Usage Mining 666
64.3 Web Servers 666
64.4 Tools 667
64.5 Literature Review 668
64.6 Conclusion 669
References 670

65 Analysis of Heart Rate Variability Using Different Techniques of Breathing 671
R. Kalpana, C. R. Sai Krishna, Nihar Jajodia, and M. N. Mamatha
65.1 Introduction 671
65.1.1 Data Acquisition 672
65.1.2 Data Conversion 673
65.1.3 Data Processing 673
65.1.4 Data Analysis 673
65.2 Linear and Nonlinear Methods 674
65.2.1 Linear Methods 674
65.2.2 Non-linear Methods 674
65.3 Results 675
65.4 T Test 676
65.5 Conclusions 676
References 680

66 Minimum Time Search Methods for Unmanned Aerial Vehicles ... 681
R. Darsini, Nikita Ganvkar, Karthik Gurunathan, and Ranjita Kumari Dash
66.1 Introduction 681
66.2 Literature Overview 682
66.3 Optimization Approaches 685
66.3.1 Ant Colony Optimization (ACO) 685
66.3.2 Evolutionary Algorithm (EA) 686
66.3.3 Cross Entropy Approaches (CEO) 686
66.3.4 Bayesian Optimization Algorithm (BOA) 686
66.4 Mathematical Evaluation of MTS 687
66.4.1 Models and MTS Evaluation In [1] 687
66.4.2 Models and MTS Evaluation In [11] 689
66.5 Conclusion and Future Work 690
References 690

67 A Comparative Study on Breast Cancer Tissues Using Conventional and Modern Machine Learning Models 693
D. Lakshmi, Srinivas Reddy Gurrela, and Manideep Kuncharam
67.1 Introduction 694
67.2 Conventional Classification Models 694
67.2.1 Data Description 694
67.2.2 Methodology 695
67.2.3 Data Standardization 695
67.2.4 Model Development 695
67.2.5 Model Improvement Techniques 695
67.3 Results and Discussion 696
67.3.1 Comparative Study on 10 Classification Models 696
67.4 Modern Method Using tSNE 698

67.5 Conclusion 699
References 699

68 Clinical Decision Support System for Knee Injuries Treatment Using Multi-Agent System 701
Naveen Dalal and Indu Chhabra
68.1 Introduction 701
68.2 Requirements of Intelligent Agents for Injury Detection and Diagnosis 702
68.3 Contribution of Fuzzy Logic in Medical Domain 703
68.4 Multi-Agent Based Diagnosis 704
68.5 Results and Performance Analysis 708
68.6 Statistical Analysis 710
68.7 Conclusion 710
References 712

69 Predictive Loan Approval Model Using Logistic Regression 715
Aiswarya Priyadarsini Behera, Siddharth Swarup Rautaray, Manjusha Pandey, and Mahendra Kumar Gourisaria
69.1 Introduction 716
69.2 Literature Survey 717
69.3 Performance of Logistic Regression and Algorithm Used 717
69.4 Experimental Setup and Result 718
69.5 Conclusion and Future Work 718
References 722

70 Efficient Load Balancing in Distributed Computing Environments with Enhanced User Priority Modeling 723
Maniza Hijab and Avula Damodaram
70.1 Introduction 723
70.2 Related Work 724
70.3 Proposed Work 724
70.4 Simulation Setup 726
70.5 Results and Discussions 727
70.6 Conclusion 729
References 729

71 Deep Neural Networks Based Object Detection for Road Safety Using YOLO-V3 731
Jalaja Tattari, Vineeth Reddy Donthi, Dheeraj Mukirala, and S. Komal Kour
71.1 Introduction 731
71.2 Related Work 732
71.3 Proposed Work 732
71.3.1 YOLO-v3 733
71.3.2 CNN 734
71.3.3 Input 735

71.3.4 Training ... 735
71.4 Experiments and Results ... 736
71.4.1 Experiments ... 736
71.4.2 Results ... 736
71.5 Conclusion and Future Work ... 736
References ... 738

72 Log-Based Authentication for Cloud Environments ... 739
Nalini Subramanian and G. Shobana
72.1 Introduction ... 739
72.2 Related Work ... 740
72.3 Proposed Framework ... 742
72.3.1 Support Vector Machine ... 743
72.4 Classification Phase Algorithm ... 743
72.5 Experimental Evaluation ... 744
72.6 Conclusion ... 745
References ... 746

73 Development of a Wireless Sensor Node for Electrical Power Distribution Monitoring System ... 747
Sanjay Chaturvedi, Arun Parakh, and H. K. Verma
73.1 Introduction ... 748
73.2 Development of Wireless Sensor Node ... 749
73.2.1 Microcontroller Selection ... 749
73.2.2 Voltage Measurement Unit ... 750
73.2.3 Current Measurement Unit ... 751
73.2.4 Frequency Measurement Unit ... 752
73.2.5 Communication Unit ... 752
73.3 Implementation of Proposed System ... 753
73.4 Conclusion and Future Scope ... 754
References ... 755

74 CommonKADS and Ontology Reasoner: Bulky-Baggage Case Study ... 757
Edwin Fabricio Lozada T., Wladimir L. Tenecota, Paul Santiago Pullas Tapia, and Livio Danilo Miniguano Miniguano
74.1 Introduction ... 757
74.2 Methodology ... 759
74.2.1 Organization Model ... 759
74.3 Experimentation ... 759
74.4 Conclusions ... 762
References ... 764

75 Accelerated Sorting of Apples Based on Machine Learning ... 765
Kiron Deb, Shobhin Basu, Vamshi Krishna Palakurthy, and Nabarun Bhattacharyya
75.1 Introduction ... 765

75.2 Description of Method ... 766
75.2.1 Training Data ... 766
75.2.2 Training Program ... 767
75.2.3 Testing Program ... 767
75.3 Results and Analysis ... 767
75.4 Validation Using Google Cloud AutoML Vision ... 769
75.5 Further Work ... 769
75.6 Conclusion ... 770
References ... 771

76 Execution of Smart Contacts Concurrently Using Fine-Grain Locking and Asynchronous Functions ... 773
Hemalatha Eedi and Pattan Asif Khan
76.1 Introduction ... 773
76.2 Related Work ... 774
76.3 Proposed Methodology ... 775
76.3.1 Implementation in Java ... 780
76.3.2 Smart Contracts Used for Results ... 780
76.4 Experimental Results ... 781
76.5 Conclusion and Future Scope ... 781
References ... 783

77 Artificial Intelligence-Based Automatic Evaluation Engine ... 785
Sai Gayatri Vadali and Gali Suresh Reddy
77.1 Introduction ... 785
77.1.1 Motivation ... 786
77.1.2 Relevant Work ... 786
77.2 Background Work ... 787
77.2.1 Spell Check ... 787
77.2.2 Grammar Check ... 787
77.2.3 Semantic Extraction ... 787
77.3 Problems Posed by Existing Systems ... 788
77.3.1 No Generic Grammar Checker ... 788
77.3.2 Extracting Context or Semantics of the Paragraph ... 788
77.3.3 Dealing with Variety of Sentences ... 788
77.3.4 Humor and Sentiment Analysis ... 789
77.3.5 Time and Space Complexity ... 789
77.4 Proposed Work ... 789
77.4.1 Preprocessing of Text ... 789
77.4.2 Syntax Analysis ... 790
77.4.3 Semantic Analysis ... 790
77.4.4 Sentiment Analysis ... 790
77.4.5 Final Score ... 791
77.5 Experimentation and Results ... 791
77.5.1 Different Types of Answers that Can Be Written by Candidate ... 791

77.6 Conclusions and Future Work 792
References 793

78 SmartRPL: Secure RPL Routing Protocol for Internet of Things Using Smart Contracts 795
P. Subhash, P. Navya Sree, and M. Pratyusha
78.1 Introduction 795
78.1.1 Routing Protocol for Low Power and Lossy Networks 796
78.1.2 Blockchain Technology 797
78.1.3 Ethereum 797
78.1.4 Smart Contract 798
78.2 Problem Statement 798
78.3 Literature Review 798
78.4 Proposed System 799
78.5 Performance Evaluation 800
78.5.1 Packet Delivery Ratio (PDR) 801
78.5.2 Throughput 801
78.6 Conclusion 802
References 803

79 Customized Music Classification and Recommendation System Based on Classifiers of Neural Networks and Sensor Embedded on Smart Devices 805
V. Roopa, A. ChristyJeba Malar, R. Rekanivetha, R. Thanga Pradeep Kumar, R. Sarveshwaran, and A. Prithiksha Parameshwari
79.1 Introduction 806
79.2 Literature Survey 806
79.3 Proposed System 808
79.4 Result and Discussion 812
79.5 Conclusion 815
References 815

80 Localization of Optic Cup-Disc in Retinal Images Using Morphological Filters for Glaucoma Detection 817
Deepika Pal, Vikrant Bhateja, Archita Johri, and Babita Pal
80.1 Introduction 817
80.2 Proposefd Segmentation Methodology Using Morphological Filters and Thresholding 818
80.3 Results and Discussions 821
80.4 Conclusion 823
References 823

81 Region Labeling Based Brain Tumor Segmentation from MR Images 825
Vikrant Bhateja, Mansi Nigam, and Anuj Singh Bhadauria
81.1 Introduction 825
81.2 Proposed Methodology for Tumor Segmentation 826
81.2.1 Region Labelling 826
81.2.2 Proposed Region Labeling Based Tumor Segmentation 828
81.3 Results and Discussions 829
81.3.1 Experimental Results 829
81.3.2 Discussions 830
81.4 Conclusion 830
References 830

82 A Comparative Evaluation of Decomposition Methods Based on Pitch Estimation of Piano Notes 833
U. Vamsi Krishna, R. Priyamvada, G. Jyothish Lal, V. Sowmya, and K. P. Soman
82.1 Introduction 833
82.1.1 Wavelet Decomposition 835
82.1.2 Variational Mode Decomposition 835
82.1.3 Dynamic Mode Decomposition 836
82.2 Dataset Description 836
82.3 Methodology 837
82.4 Results and Discussions 837
82.4.1 Selection of Metrics 837
82.4.2 Parameter Tuning 839
82.5 Conclusions and Future Scope 841
References 842

Author Index 845

About the Editors

Suresh Chandra Satapathy is currently working as Professor, KIIT Deemed to be University, Odisha, India. He obtained his Ph.D. in Computer Science Engineering from JNTUH, Hyderabad, and master's degree in Computer Science and Engineering from National Institute of Technology (NIT), Rourkela, Odisha. He has more than 27 years of teaching and research experience. His research interest includes machine learning, data mining, swarm intelligence studies and their applications to engineering. He has more than 98 publications to his credit in various reputed international journals and conference proceedings. He has edited many volumes from Springer AISC, LNEE, SIST and LNCS in past, and he is also the editorial board member in few international journals. He is a senior member of IEEE and a life member of Computer Society of India. Currently, he is National Chairman of Division-V (Education and Research) of Computer Society of India.

Vikrant Bhateja is Associate Professor, Department of ECE in SRMGPC, Lucknow. His areas of research include digital image and video processing, computer vision, medical imaging, machine learning, pattern analysis and recognition. He has around 160 quality publications in various international journals and conference proceedings. He is associate editor of IJSE and IJACI. He has edited more than 30 volumes of conference proceedings with Springer Nature and is presently EiC of IGI Global: IJNCR journal.

Dr. Margarita N. Favorskaya is Professor and Head of the Department of Informatics and Computer Techniques at Reshetnev Siberian State University of Science and Technology, Russian Federation. Professor Favorskaya is a member of KES organization since 2010, the IPC member and Chair of invited sessions of over 30 international conferences. She serves as a reviewer in international journals (Neurocomputing, Knowledge Engineering and Soft Data Paradigms, Pattern Recognition Letters, Engineering Applications of Artificial Intelligence), an associate editor

of Intelligent Decision Technologies Journal, International Journal of Knowledge-Based and Intelligent Engineering Systems, International Journal of Reasoning-based Intelligent Systems, a honorary editor of the International Journal of Knowledge Engineering and Soft Data Paradigms, reviewer, a guest editor, and a book editor (Springer). She is the author or the co-author of 200 publications and 20 educational manuals in computer science. She co-authored/co-edited seven books for Springer recently. She supervised nine Ph.D. candidates and presently supervising four Ph.D. students. Her main research interests are digital image and videos processing, remote sensing, pattern recognition, fractal image processing, artificial intelligence, and information technologies.

Dr. T. Adilakshmi is currently working as Professor and Head of the Department, Vasavi College of Engineering. She completed her Bachelor of Engineering from Vasavi College of Engineering, Osmania University in the year 1986 and did her Master of Technology in CSE from Manipal Institute of Technology, Mangalore, in 1993. She received Ph.D. from Hyderabad Central University (HCU) in 2006 in the area of Artificial Intelligence. Her research interests include data mining, image processing, artificial intelligence, machine learning, computer networks and cloud computing. She has 23 journal publications to her credit and presented 28 papers at international and national conferences. She has been recognized as a research supervisor by Osmania University (OU) and Jawaharlal Nehru Technological University (JNTU). Two research scholars were awarded Ph.D. under her supervision, and she is currently supervising 11 Ph.D. students.

Chapter 1
Pad Vending Machine with Cashless Payment



Anjali Prajapati and Anandu M. Dharan

Abstract A vending machine aims to provide required product or the service to the customer with certain ease, wherein not much effort is required. This research work aims to design a pad vending machine with an option of payment using QR code which is implemented using blockchain to make the system much more efficient and reliable than the existing systems present in the Indian market. The system is divided into two parts, first being the working of the machine and second being the mode of payment which is implemented using a blockchain. It is noticed many times that due to unpredictable menstrual flow women tend to face a lot problems. To overcome this problem, a pad vending machine is proposed with certain advancements through which women can help themselves in the stated circumstances.

1.1 Introduction

There are different types of networking systems through which a technology or an application can be implemented. There are majorly three types of networking systems which are majorly used to implement any application and are classified into three, namely centralized, distributed, and decentralized. Each type of networking system serves in a different and unique way to address the problem, and almost every application can be implemented using each of the network system; however, we should always look for the most efficient way of implementing any application in a network. In the current scenario, most of our applications are implemented on a centralized network which follows a three-tier architecture which includes a server, a client, and a receiver. Since there are three layers involved in the architecture, the process consumes a lot of time in sending and receiving requests. But if the server which is present in the middle of the receiver and the sender is removed, it would save a lot of time in the networking process by making the whole process a decentralized one, which makes the process easy as it has only two parties which are

A. Prajapati (✉) · A. M. Dharan
Christ (Deemed to be University), Faculty of Engineering, Bangalore, India
e-mail: anjali.prajapati@btech.christuniversity.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_1

communicating to each other directly. Everything around us which is implemented through centralized system can be switched to a decentralized one. For example, a simple money transaction; consider A wants to send Rs. 100 to B; in a centralized network, A will send a request to the bank for the transaction of Rs. 100, and once the bank approves the transaction, the money would be credited to B's account. Consider the same situation in a decentralized network; since there is no middle layer involved, there would not be any requirement of the approval from the bank, A can directly transfer the money to B's account, and the bank would only record that on a cloud or in a database. Hence, a decentralized system is introduced to make any process simpler, easier and hassle-free with full security.

1.2 Literature Review

The concept of vending machine is to provide the required product to the customer through a machine with just a push button. There are different vending machines available in the market, for example coffee vending machine, food and beverages vending machine, pad vending machine, etc. Most of these machines available in the Indian market requires their customers to have hard cash for the payment to be done. However, there are some food and beverages vending machine available in the Indian market which provides an option of payment using Paytm. It is majorly required and essential as not every time the customer would have a hard cash to do the payment and purchase the product.

In [1], the author has defined vending machine as a terminal device that interacts with the ultimate customer for the direct selling the proposed product or service and aims to design a system that makes the customer service faster and efficient. The author also aims to reduce the maintenance cost of the whole system and making it more interactive.

In [2], the author informs the readers about the success rate of vending machine in Japan and USA because of cashless payment and ease of access to the products required. It also helped these countries in reducing the crime rates because of cashless payment methods.

In [3], the author puts up a thought about the security of the payment using mobile phones, and it requires authentication as it can be hacked. The payment can be done using mobile phones through QR code. QR code can store more information than a traditional barcode, and the security of the system can be maintained by using Public Key Infrastructure (PKI) which adds an additional layer of security to the system.

A lot of research has been made to make sure any vending machine available in the market should be interactive enough for the customers and should have fast and secure mode of payment with an ease of use. Specifically, in the case of a pad vending machines which are used by women in case of emergency of unpredicted menstrual flow. There are different researcher who worked on the different parts of the vending machine to make it efficient and fast for its customers. The major problems faced by the customer while using the vending machine was the machine

being less interactive and the mode of payment. As said earlier, majorly the mode of payment in a vending machine is via coin system or cash payment, and it is not necessary that every time a user would have a coin or cash; hence, an addition of smart and easy mode of payment would really help the customers, especially in case of a pad vending machine.

1.3 Proposed System

The proposed system is a pad vending machine which has an interactive UI for the customer, and the customers can scan the QR code at the end of the process for completing the payment process. An architecture of the model is given below (Fig. 1.1):

The proposed system aims to develop a decentralized system using blockchain and remove the third-party involvement; hence, the system has no administration. The architecture of the Pad Vending Machine (PVM) includes only two parties, namely the user and the vendor and two status in two different variable stack.

The description of each and every and every part of the system is given below**:**

A. **User X**: A random customer *X* would come to the PVM and will select the size of the sanitary napkin required via a push button; the system would check the availability of the same, and if it is available, it will be delivered to the customer once the payment is done. In case, the select size of the napkin is not available, the user would be asked by the system to opt another size of the pad.
B. **Vendor**: The second involvement in the system would be of a vendor who would be responsible for refilling the PVM with the napkins of different sizes as and when required, and the payment of each napkins would be done by the

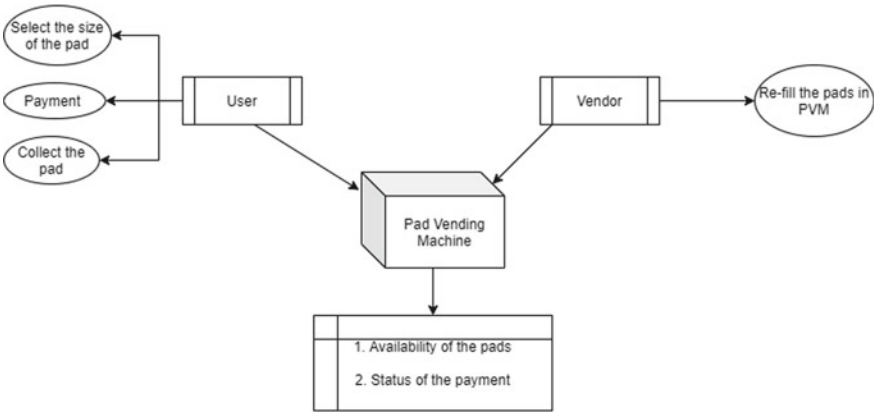


Fig. 1.1 Architecture of the model

customers via QR which would be directly send to the vendor's account without any approval from third party as it uses blockchain for the payment to be done.

C. **Availability of the pad**: In PVM, there is a stack which keeps a check on the availability of the sanitary pads of each size. For each size of the sanitary napkin, there are two values to be updated, and those are 0 and 1. '1' defines that there are napkins available in the machine, and '0' defines that there are no napkins left in the machine of that size.

D. **Status of the payment**: It is a single variable which is defined to check the payment status, and it contains only Boolean values 0 and 1. '1' defines the payment is done, and '0' defines the payment is pending.

1.4 Implementation

The sequence of the steps involved in the proposed system is shown in the flow chart given in Fig. 1.2. The proposed system is a pad vending machine which uses QR scanning method for payment which is implemented using blockchain. The flow of the events in the system is quite simple and easy. The process starts with a user standing in front of the machine in the need of the sanitary napkin. In the first step, the user is expected to use the push button present in the machine to select the size of the required sanitary napkin.

The second step is to scan the QR code available on the screen of the machine, through which the payment is supposed to be done. Since the mode of payment is through a QR code and it is implemented using a blockchain mechanism, each time the user takes a pad, the machine generates a new QR code for the payment to be done to the same vendor, which ensures the safety of the system as it is anonymous and cannot be hacked. The mode of the payment is so designed using blockchain which makes the system decentralized wherein the system is governed by a network of nodes and all nodes works according to the predefined protocol. Also, all the funds transfer are controlled by the users and not a centralized party like a bank.

As soon as the machine generates a new QR code to be scanned, it actually makes a request for the transaction to be done. The transaction is represented as a 'block.' Thereafter, the block is broadcasted to every party in the network. Once the network of nodes receives the transaction, they validate it. Now, a new block of data is added to the existing blockchain and the transaction is completed. The existing blockchain is a data structure that keeps the record of all the transaction made to the vendor, and a new block is added every time a new payment is done. The transaction would be completed in less than a minute as the transaction does not required any sort of approval from the bank.

Once the payment is done by the user and the vendor receives it, the payment variable would be turned to '1' from '0' automatically stating that the payment is done, and now, the sanitary napkin can be delivered to the user, and thereafter, the sanitary napkin would be delivered.

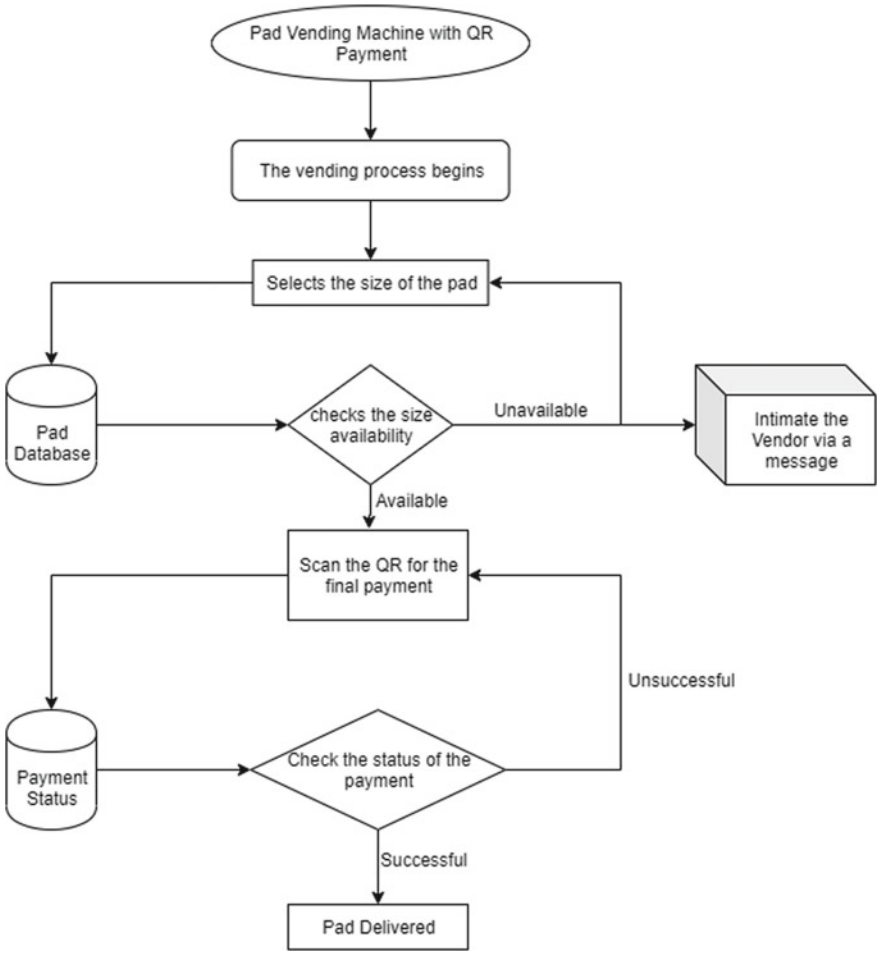


Fig. 1.2 Flow chart of the process

On the other hand, if a situation arises wherein the selected size of the pad is not available, the value in the pad availability stack for that particular size of the pad would be turned to '0′ and a message would be sent to the vendor immediately for the re-fill of the pads.

1.5 Results and Comparative Study

The system provides the sanitary napkin of the required size on the basis of the customer's request. Once the request is made to the machine, it checks for the availability of the size of the napkin and generates a QR code for the payment to be done

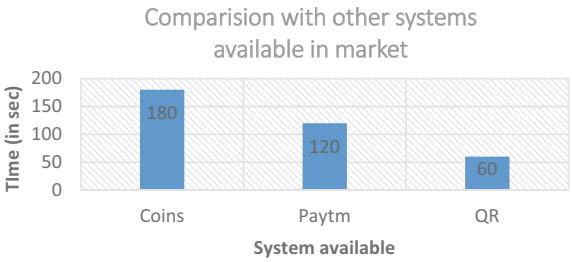


Fig. 1.3 Comparative study of the proposed system with existing systems

to the vendor. As soon as the payment is done and verified by the system, the sanitary napkin is delivered to the user. However, if there a situation wherein the selected size of the napkin is not available, the machine asks the user to select a different size and procedure with it.

The system is compared with different existing systems present in the Indian market and stands out different because of its fast processing and efficient working, and a graphical representation of the same is shown below in Fig. 1.3. With the introduction of blockchain in the payment, the system becomes really efficient and secure, uses less time to process the request and the density of providing the service increases.

1.6 Future Scope

The stack which is being used for storing the information about the availability of the size of the sanitary napkins and the variable which stores the information regarding the status of the payment can be replaced with two other blockchain structures. This would again reduce the cost of the whole system as there would not be any requirement of storage. Also, the efficiency of the system would increase exponentially.

1.7 Conclusion

The system proposed above is a pad vending machine in which the mode of payment is the major advancement as it is implemented using a blockchain mechanism which makes the system faster in terms of processing the request and serving the customer in less than a minute. Also, it makes the system a decentralized peer-to-peer network, and if the data is entered once, it cannot be altered by any other unreliable source.

References

1. Rifa, H., Budi, R.: Blockchain Based E-Voting Recording System Design. IEEE (2017). 978-1-5386-3546-9/17/$31.00
2. Koç, A.K., Yavuz, E., Çabuk, U.C., Dalklç, G.: Towards Secure E-Voting Using Ethereum Blockchain. IEEE (2018) 978-1-5386-3449-3/18/$31.00
3. Kshetri, N., Voas, J.: Blockchain-Enabled E-Voting. IEEE (2018). 0740-7459/18/$33.00
4. Christian.: Desain Dan Implementasi Visual Cryptography Pada Sistem E-Voting Untuk Meningkatkan Anonymity. InstituteTeknologi Bandung (2017)
5. Dougherty, C.: Vote Chain: Secure Democratic Voting (2016)

Chapter 2
Design of Wrung Order Layout Using Advanced Optimization Techniques with Integrated Variable Batch Scheduling



K. Mallikarjuna and Y. Hariprasad Reddy

Abstract Universally, specialists and scientists accept that flexibility assumes a elementary play in modern factory segment. Only associated with modest parcel size generation since agility adaptable is an indispensable part to be incorporate into course of action of racks in format plan among the assembling fragment. In view of such conditions, considering NP hard double target issues is, regularly, a lumbering responsibility. In this work, researchers tended to about a populace-based high end search techniques like differential development (DE) and sheep run technique (SRT) for making wrung structure configuration matters in lithe system of manufacturing location. The instigators focused on twofold aim headway connected with fundamental objective is stressed over the versatile slot (FJSP) arranging issue, the accompanying objective focused on wrung order layout matters where expelling the interest of machineries within lead-ins of wrung steps to control rigid transference cost and hoarding working time of employments on machineries. The accomplishment of the estimation (SRT and TS) is crisscross by standard issues. At long last, it is pondered that SRT yields better outcomes at the point on par with TS.

2.1 Introduction

In the present situation, mechanized assembling ventures are under prodigious stress which brought about by the increasing expense of vitality, materials, works, capital, and strengthening overall challenge. While these patterns will stay for quite a while, the issue fronting producing today runs much cavernous. By and large, they come

K. Mallikarjuna (✉)
Department of Mechanical Engineering, G. Pullaiah College of Engineering and Technology, Kurnool, India

Y. Hariprasad Reddy
Rayalaseema University College of Engineering, Kurnool, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_2

from the very idea of the assembling procedure itself. So as to beat that, adaptable manufacturing frameworks (@FMSs) are viewed as one of the most productive strategies to use in lessening or taking out assembling issues. FMS is in excess of a specialized arrangement [1]; it is a factory-based driven arrangement prompting to improve gainfulness via decreasing process times and stock levels and improved assembling viability through expanded operational adaptability, consistency, and control.

The FMS design includes assigning different hold for achieving full skill. The plan has an impact eager for elapsed time and cost [2] which ought to be resolved in the beginning of the FMS [3]. By and by, the most generally utilized sort of FMS designs [4] is as per the following:

1. Line or single column design
2. Circle format or oval shape design
3. Stepladder or pecking order or wrung structure format
4. U-shaped format.

Amongst the above designs, this task addressed about wrung order layout with scheduling as restrictions utilizing sheep run technique (SRT) and Tabu search method (TS).

2.2 Problematic Depiction

The work plan got from Abraham et al. [5] cited for this work. Analyst's complement on construction of wrung position in versatile course of action of collecting with [FJSP] versatile job shop booking as impediment.

2.2.1 Multi-goal Mathematical Representations

Here, maker familiarizes the twofold aim condition with the versatile variable bunch arranging problems which are united with wrung design arrangement primes to lessen the goal of work such as throughput, make run,

minimize MAKSP, $F\ (S_{\text{maxi}})$

$$\text{Minimize, } F(S_{\max}) = S_{n,m} \tag{2.1}$$

Connected to

$$S_{i,j,k} \leq S_{i,j+1,k} - T_{i,j+1}, \quad \text{for} \quad j \ \ldots \ = \ 1,\ 2,\ 3 \ \ldots \ p$$

$$S_{i,j,k} \geq 0, \quad \text{for} \quad j \ \ldots \ = \ 1,\ 2,\ 3 \ \ldots \ p$$

2.3 Proposed Advanced Optimization

This work is proposed with two advanced non-traditional optimization techniques such as

(a) Sheep run technique
(b) Tabu search method

The above two methods are applied on to multiobjective function of wrung order layout design with scheduling constraints, and necessary throughputs, make run have been obtained.

2.3.1 Sheep Run Technique

In the inception, sheep run techniques (SRT) was developed by Hyunchul and Byungchul [6]. This calculation was started to explain huge scale issues on planning over a period of a few continuous years. And it is competing with performance of another evolutionary algorithm called genetic algorithm calculation, and it is alluded as staggered hereditary activities can get great arrangements. This calculation was commonly founded on the common difference in sheep in the group.

Let us consider two runs of sheep in a structure with two persons for vigilance who being monitoring and observing those sheep runs. At the moment when these two were busy in chitchat [7], there is a probability of unification of herds been taken place, and in such condition, both the persons arrive at the mixed gathering and endeavor to segregate the sheep from mixed surges and keep the groups as of now [8]. In actuality, the effect of inheritance will cause imperfection on specific surges. In this, necessary steps which describe the multilevel genetic operation for scheduling problem which is a constraint in layout design in sheep run techniques are given below.

Multilevel Genetic Operation

```
int b1Makespan, b2Makespan;
            int b1=1, b2=2;
            b1Makespan = chsMakespan[1];
            b2Makespan = chsMakespan[2];

            //Best 2 chromosomes
            for (int p = 2; p <= no_populations; ++p) {
                  if(b1Makespan > chsMakespan[p]){
                        b1Makespan = chsMakespan[p];
                   }     b1 = p;
             }

cross overing the strings using subchromosomal
            for (inte p = 1; p <= no__ populations;++ p)
{
                  inte *ttt;
                  ttt = subbChromosomeCrossOvers(chss[p],
sizee*subChssCrossOverRatios, size);
                  for (inte b = 1; b <= sizee; b ++ ) {
                        chsSUB[pp][b] == ttt[b];
                  }
            }
```

2.3.2 Tabu Search

Tabu search (TS) is a heuristic strategy initially proposed by Glover in 1986 to different combination problems have noted up in the activities explore writing. In a few cases, the strategies portrayed give arrangements near optimality [9] and are among the best, if not the best, to handle the troublesome issues within reach. These victories have made TS incredibly well known among those keen on discovering great answers for the huge combinatorial issues experienced in numerous down to earth settings. Here, a necessary procedural step for scheduling using Tabu search [10] is furnished through pseudocode.

Pseudocode

```
int
makespan_OdjectWithWaitingTimes(Schedulable_operations
representatives1[],int sequence[]){
            Schedulable_operations
S[MAX],J[MAX],Job_order[MAX],RR[MAX],JJ[MAX];
            int
R[no_machines+1],RT[no_machines+1],SMPT[no_batchs+1],OP[n
o_batchs+1],
            nn=no_batchs*no_operations,order=1,n=0;
            int num=0;
            int OP1[MAX];
            int Morder[MAX_MC][MAX];
            int Morder1[MAX_MC][MAX];
            initialize();

      //display_RepresentativesSequence(representatives1,se
quence);
            for(int k=1;k<=no_machines;k++){
                  for (int l = 1; l<=
no_batchs*no_operations; l++) {
                        Morder[k][l] = 0;
                        Morder1[k][l] = 0;
                  }
                  OP1[k] = 1;
            }
```

2.4 Results and Discussion

Six different machine slots with different processing of jobs on machines to be allocated in machine slots in wrung order layout for effective design of layout are considered, along with necessary process times, interslot distance between machines and load and unloading cost is also considered. Different combinations of these six slots for machine allocations, six jobs, and varied operations of jobs for different cluster combinations were used, and four layouts are used to generate five example problems [1]. Here, Table 2.1 shows production summary of wrung order layout, and Table 2.2 shows batch variabilities with batch magnitudes of wrung order layout.

Table 2.1 Framework of manufacture system

Layout pattern	Number of machines	Number of batches	Number of processes	Load/unload stations
Wrung order	6	6	6	2

Table 2.2 Batch variabilities with batch magnitudes of the pecking layout with six machines with six batches

Batch No.		B1	B2	B3	B4	B5	B6
Batch variabilities	VBS	100	70	55	60	80	40

Table 2.3 Assessment of make run of the proposed evolutionary algorithms (for VBS with 100 generations)

Problem Instance	SFHA	TS
	MAKRN (min)	MAKRN (min)
KMK 1—(6 × 6 × 6)	5020	5520
KMK 2—(6 × 5 × 5)	4200	4800
KMK 3—(6 × 5 × 4)	3445	3950
KMK 4—(6 × 4 × 4)	3025	3300
KMK 5—(6 × 4 × 3)	2554	2995

Table 2.4 Assessment of overall conveyance cost of the projected evolutionary algorithms (VBS through 100 generations)

Problem Instance	SFHA	TS
	OCC	OCC
KMK 1—(6 × 6 × 6)	8400	9200
KMK 2—(6 × 5 × 5)	7652	8560
KMK 3—(6 × 5 × 4)	7025	8020
KMK 4—(6 × 4 × 4)	6654	7885
KMK 5—(6 × 4 × 3)	5951	6995

2.4.1 Inferences

By thorough analysis From Tables 2.3 and 2.4 and Figs. 2.1 and 2.2 it is revealed that accomplishment of SRT and TS for figuring Overall Conveyance penalty and Makerun (MAKRN) is weakening as issue ration is littler according to the issue scope. By relative investigation, it is seen that OCC are progressed for SRT.

2.5 Conclusion

This work addressed on the presenting of wrung order layout configuration with incorporated planning for which the recurrence of excursions of jobs amid machines, the gap among the machines with stacking and emptying good ways from stacking/emptying place to all machineries, and component material taking care of

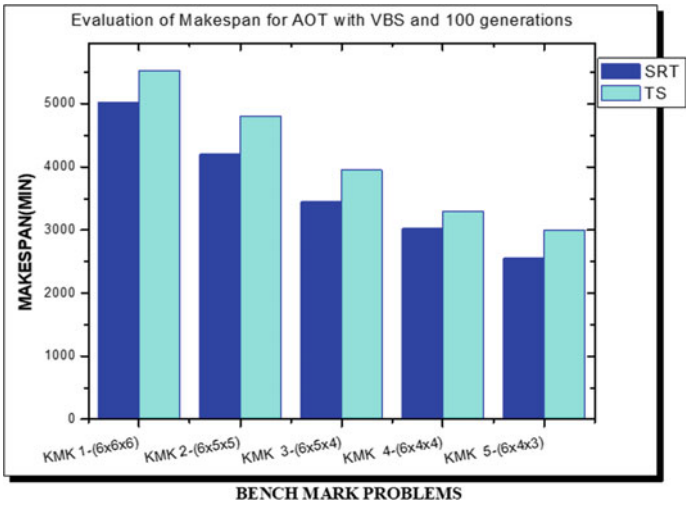


Fig. 2.1 Plot of make run to evaluate performance of projected algorithm

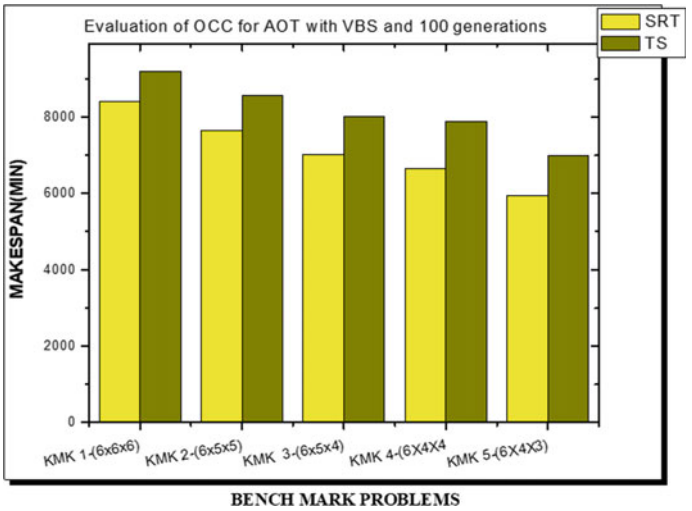


Fig. 2.2 Plot of overall conveyance charge to evaluate performance of projected algorithm

expense (MHD) are evaluated in an unexpected way. The issue is encircled as the quadratic project issues (QPI) proposal of capacity layout problems. This is inferable from the point that in the QPI models, the separation between the places of allocation is recognized well ahead of time yet it is structure subordinate for the problems considered in this paper. From the outcomes, we infer that wrung order design is upgraded utilizing SRT and is superior to TS with varible batch sizes and constant MHD cost and recurrence of outings between machines.

References

1. Mallikarjuna, K., Veeranna, V., Reddy, K.H.: A new meta-heuristics for optimum design of loop layout in flexible manufacturing system with integrated scheduling. Int. J. Adv. Manuf. Technol. **84**(9–12), 1841–1860 (2016). https://doi.org/10.1007/s00170-015-7715-9
2. Apple, J.M.: Plant Layout and Material Handling, 3rd edn. The Ronald Press Co, New York (1977)
3. Satheesh Kumar, R.M., Asokan, P., Kumanan, S.: Design of loop layout in flexible manufacturing system using non-traditional optimization technique. Int. J. Adv. Manuf. Technol. **38**(5–6), 594–599 (2008). https://doi.org/10.1007/s00170-007-1032-x
4. Groover, M.P.: Production Systems and Computer Integrated Manufacturing. Automation Prentice-Hall Inc. (1992). ISBN 0-87692-618-9
5. Liu, H., Abraham, A., Grosan, C., Li, N.: ICDIM 2007. LNCS, vol. 1, pp 138–145. https://www.springer.com/lncs. Accessed 21 Nov 2016
6. Kim, H., Ahn, B.: A new evolutionary algorithm based on sheep flocks heredity model. IEEE Pacific RIM Conf. Commun. Comput. Signal Process. Proc. (2001). https://doi.org/10.1109/pacrim.2001.953683
7. Subbaiah, K.V., Nageswara Rao, M., Narayana Rao, K.:. Scheduling of AGVs and machines in FMS with makespan criteria using sheep flock heredity algorithm. Int. J. Phys. Sci. (2009)
8. Nara, K., Takeyama, T., Kim, H.: New evolutionary algorithm based on sheep flocks heredity model and its application to scheduling problem. Proc. IEEE Int. Conf. Syst. Man Cybernet. (1999). https://doi.org/10.1109/icsmc.1999.816603
9. Glover, F.: Tabu search—Part I. ORSA J. Comput. (1989). https://doi.org/10.1287/ijoc.1.3.190
10. Ekşioğlu, B., Ekşioğlu, S.D., Jain, P.: A tabu search algorithm for the flowshop scheduling problem with changing neighborhoods. Comput. Ind. Eng. (2008). https://doi.org/10.1016/j.cie.2007.04.004

Chapter 3
Adaptive Beamforming Using Radial Basis Function



Gowtham Narayan, Sharath Kumar, Abhinav Bharadwaj, Srikar Vangala, and R. Lavanya

Abstract With the advancements in 5G mobile communication, the need for improvements in spatial signal processing has become indispensable. Beamforming is one such technique used to transmit or receive wireless signals in a particular direction in space and suppress interfering signals. This is achieved by making use of a uniform antenna array. This paper focuses on the receiver antenna. MUSIC algorithm is highly popular for calculating the direction of the received signal. However, it is computationally expensive and hence not suitable for real-time applications. The idea proposed in this paper is to develop a neural network model based on MUSIC, to cater to real-time need of modern communication systems.

3.1 Introduction

Antenna arrays provide an effective way to track and process signals arriving in various directions in wireless communication. An array of sensors can arbitrarily alter its beampattern by altering the weights of its array elements compared to a single antenna that is constrained in directionality and bandwidth [1]. Ideally, maximum signal reception has to be towards the direction of arrival (DoA) of the desired signal, and minimal signal reception has to be towards the directions of the interfering signals. Among the algorithms for DoA estimation, MUSIC provides more accurate results at the cost of comparatively higher computational power [2, 3] (Fig. 3.1).

In real-world scenarios, the desired and interfering signals change their directions continuously. A fast beamforming technique is required to keep track of these changes in directions. Though the MUSIC algorithm is known for its accuracy, the computational complexity makes it unsuitable for real-time applications such as communication-on-the-move applications. To alleviate this problem, a neural

G. Narayan · S. Kumar (✉) · A. Bharadwaj · S. Vangala · R. Lavanya
Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

R. Lavanya
e-mail: r_lavanya@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_3

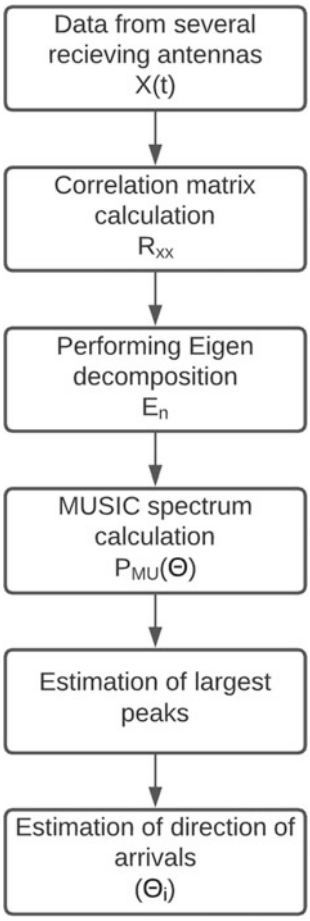


Fig. 3.1 MUSIC algorithm flowchart

network is employed to learn the function underlying the MUSIC algorithm such that the latter's accuracy is retained and at the same time the processing speed is high. Neural networks are very efficient in pattern recognition and function approximation and are becoming increasingly popular in diverse areas and applications [4, 5]. A neural network with radial basis as the activation function is employed in this work. The network output is a linear combination of inputs' radial base functions and neuron parameters [6].

3.2 Methodology

The first step for developing the neural network model is to generate dataset with received signals as input and DoA as output. This is done by varying the angle between

[+90º −90º] and sampling the outputs from each element. The signal received by the ULA with N elements is modelled as

$$X(t) = A(\theta)S(t) + N(t) \tag{3.1}$$

where

$$X(t) = \left[x_1(t)\ x_2(t)\ \ldots\ x_N(t) \right]^{\mathrm{T}} \tag{3.2}$$

$$N(t) = \left[n_1(t)\ n_2(t)\ \ldots\ n_N(t) \right]^{\mathrm{T}} \tag{3.3}$$

and

$$S(t) = \left[s_1(t)\ s_2(t)\ \ldots\ s_M(t) \right]^{\mathrm{T}} \tag{3.4}$$

$X(t)$ is a complex vector of antenna array outputs. $N(t)$ is the noise vector caused by the channel and is modelled as White Gaussian noise. $S(t)$ is the actual source signals vector transmitted with M individual signals. The $N \times M$ matrix A of the incident source signals can be defined as

$$A(\theta) = \left[a_1(\theta)\ a_2(\theta)\ \ldots\ a_M(\theta) \right]^{\mathrm{T}} \tag{3.5}$$

where

$$a(\theta_m) = \left[\mathrm{e}^{-j(n-1)\Psi_m} \right]^{\mathrm{T}} :\ 1 \leq n \leq N \tag{3.6}$$

Ψ_{m} represents the phase shift from element to element along the array which is defined by

$$\Psi_m = 2\pi \left(\frac{d}{\lambda} \right) \sin(\theta_m) \tag{3.7}$$

where d is the element spacing and λ is the wavelength of the received signal. After varying the DoA and generating the dataset for the neural network, the dataset is split in the ratio of 7:3. 70% of the dataset is used for training with multiple epochs till $R = 1$ is achieved. The remaining 30% of the dataset is used to test the neural network model.

The neural network architecture chosen was RBFNN due to its short learning time, easy design and strong tolerance to noise. RBFNN uses the activation functions of radial basis functions. The input to the neural network is the spatial covariance matrix R of signals received by the elements of the ULA.

The spatial covariance matrix is given by

$$
\begin{aligned}
R &= E\left\{X(t)X(t)^H\right\} \\
&= \mathrm{AE}\left\{S(t)S(t)^H\right\} + E\left\{N(t)N(t)^H\right\} \\
&= \mathrm{APA}^H + \sigma^2 I \\
&= \sum_{i=1}^{M} \lambda_i e_i e_i^H
\end{aligned} \tag{3.8}
$$

P denotes the power matrix of the obtained signals by the array antenna. $\underline{\lambda_i}$ denotes the eigenvalues of matrix R, σ^2 is the variance of independent White noise received by the array antenna.

This spatial covariance matrix is used as input to MUSIC algorithm to obtain DoA's. The spatial covariance matrix along with the DoAs is used as ground truths for the neural network.

3.3 RBFNN Overview

A three-layer feed forward neural network is a radial basis function (RBF). Radial base function networks are commonly used for approximation problems as a type of artificial neural network. They are differentiated from other neural network methods because of their ability to perform universal approximation and have a faster learning speed. As stated earlier, the input layer, hidden layer and output layer are the three layers that are associated with this particular neural network system. They have their own set of tasks to be performed. A brief explanation on the RBF network is presented as follows. The process of RBF training is completed once the computed error reaches the desired values or the total number of iterations of training is terminated. Commonly, an RBF network is chosen based on a specific number of nodes in the hidden layer. Computational units for transfer functions are generally handled by a Gaussian function. In comparison with multi-layer perceptron, the amount of time needed for completion of training is usually less for RBF. The neural network of radial basis function (RBFNN) consists of many elementary processing units called neurons, organised into layers. The subsequent adjacent layer is related to each neuron in a layer. There are no connections between neurons of the same layer. Neurons connect with each other by weight, and according to the received signals, they can be triggered accordingly. There are three layers in the RBFNN, namely input, output and one hidden layer, which are interconnected in a feed forward manner. The mapping is nonlinear from the input to the hidden layer, while the mapping from the hidden layer to the output layer is linear (Fig. 3.2).

The mapping functionality relies on the distance between the vector of the input and the vector of the centre. The RBFNN with input x of dimension N and output y of dimension M can be expressed as a weighted sum of a finite number of radial basis functions given by

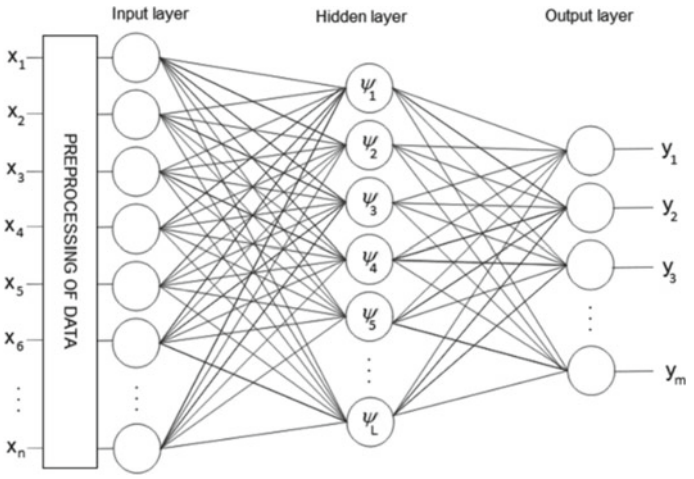


Fig. 3.2 Architecture of RBFNN

$$y = F(x) = \sum_{i=1}^{L} w_i \Psi(\|x - x_i\|) \tag{3.9}$$

$\Psi(\|x - x_i\|)$ is the radial basis function calculated by shifting $\Psi(x_i)$. L is the set of arbitrary functions and x_i is the centre of RBF $\Psi(\|x - x_i\|)$. $\Psi_s(x_i)$ is given by

$$\boldsymbol{\Psi}\left(\boldsymbol{e}^{-x^2}\right) \tag{3.10}$$

RBF is a strongly nonlinear function which offers optimal functionality for progressive learning. As a transition function, the neurons in the secret layer of RBFNN have RBF. By integrating the weighted inputs and biases, RBF determines its weighted inputs with the Euclidean distance weight function and obtains its net input. During the training phase, neurons are eventually introduced to the secret layer.

Many learning strategies exist for training RBFNN. One technique is to use the K-means algorithm to classify the initial centres of the hidden layer's Gaussian functions. The standard deviation of the Gaussian function is the mean distance of the other Gaussian functions to the first hundred closest neighbours. This method allows the weights from the input to the hidden layer to be defined. The weights from the hidden layer to the output layer, implemented on single-layer networks, are calculated by supervised learning known as the delta law.

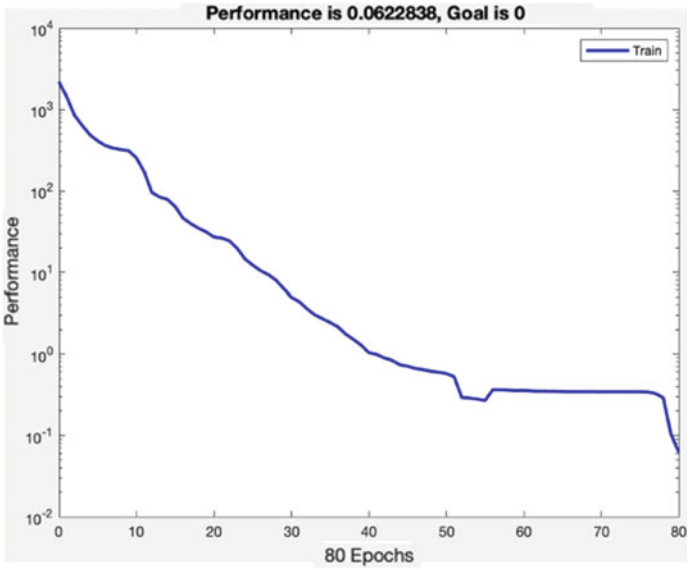


Fig. 3.3 MSE plot

3.4 Results

A 10 element ULA with helical antenna elements and element spacing given by $d = \lambda/2$ is considered for the study. The input to the neural network is given as the spatial correlation matrix of the array output vector. Before feeding the spatial correlation matrix to the neural network, pre-processing of the data has to be performed. When the matrix is separated into its real and imaginary parts, they are both symmetrical matrices. Also, the diagonal elements of the imaginary matrix are all zeros. Thus, only the elements of the upper triangular matrix are taken as input to the neural network to decrease the complexity, thus providing us with $N \times (N-1)$ inputs rather than Nˆ2 inputs. Training, validation and testing were performed on the developed neural network model. The output of the neural network is the DoA of the incoming signals. The mean squared error (MSE) of the developed neural network model was found to be 0.0622838 (Figs. 3.3, 3.4, 3.5 and 3.6).

3.5 Conclusion

In this work, a neural network-based DoA estimation algorithm is proposed. The proposed scheme employs radial basis network which is trained based on MUSIC algorithm, combining the merits of fast convergence of RBFNN and high performance of MUSIC. The algorithm is highly suitable for deployment in real-time

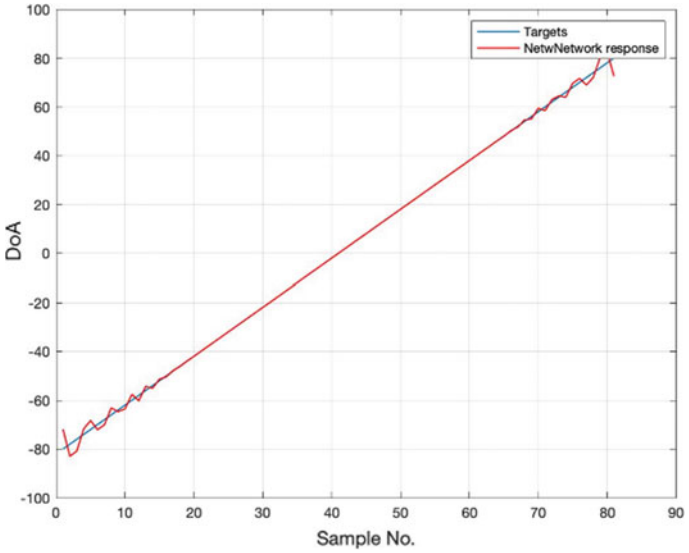
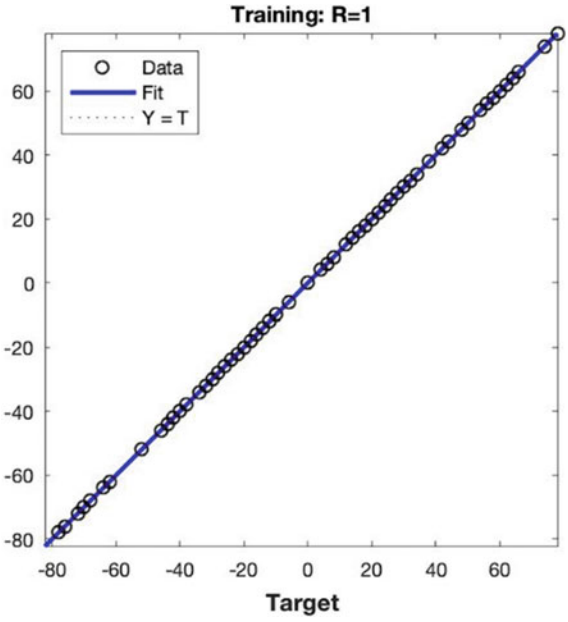


Fig. 3.4 Comparison of target DoA and RBFNN-estimated DoA

Fig. 3.5 Training regression plot



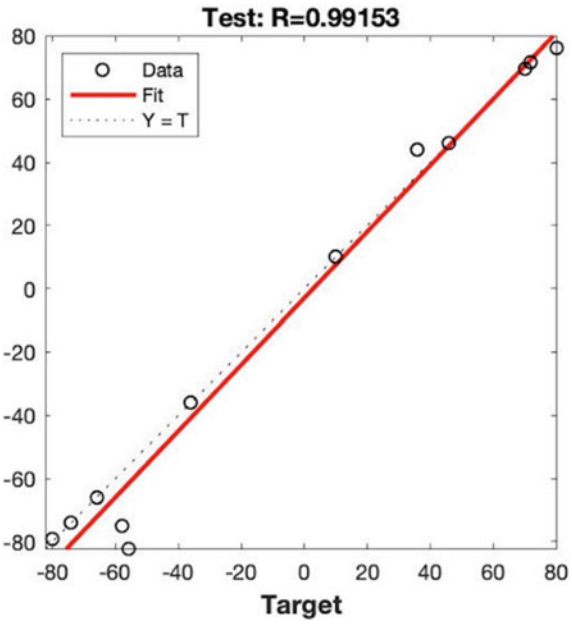


Fig. 3.6 Testing regression plot

communication applications. The work can be further extended for beamforming module as well.

References

1. Pei, B., Han, H., Sheng, Y., Qiu, B.: Research on smart antenna beamforming by generalized regression neural network. In: 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013), pp. 1–4. KunMing (2013)
2. Vikas, B., Vakula, D.: Performance comparision of MUSIC and ESPRIT algorithms in presence of coherent signals for DoA estimation. In: International conference of Electronics, Communication and Aerospace Technology (ICECA) Coimbatore, pp. 403–405 (2017)
3. Qian, R., Sellathurai, M., Wilcox, D.: A study on MVDR Beamforming applied to an ESPAR antenna. IEEE Signal Process. Lett. **22**(1), 67–70 (2015)
4. Praveena, S., Lavanya, R,: Superpixel based segmentation for multilesion detection in diabetic retinopathy. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 314–319. Tirunelveli, India (2019). https://doi.org/10.1109/ICOEI.2019.8862636
5. Rajendiran, G., Kumar, C., Vishnuprasad K., Ramachandran, K.I.: Feature mapping techniques for improving the performance of fault diagnosis of synchronous generator. Int. J. Prognost. Health Manage. **6**, 1–12 (2015)
6. Ping, Z.: DOA Estimation method based on neural network. In: 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), pp. 828–831. Krakow (2015)

Chapter 4
A Cloud-Based Centralized Smart Electricity Billing and Payment Management System



P. Pramod Kumar, Nethaji Achha, G. Ranadheer Reddy, and V. Pranathi

Abstract Energy sources are very important in our day-to-day life. One of the most important energy sources is electricity; we pay for the electricity which we use. As we are living in and around the space of artificial intelligence, we need to develop ourselves to live with artificial intelligence. In past five years, there are many changes took place in the electricity department to provide better services to the consumers. Many of the people suspect that many malpractices are going on in the electricity department while issuing a bill for consumed units and at payment time. To avoid this type of malpractices, we can go with the atomized billing system with artificial intelligence. This artificial intelligence can reduce the usage of paper and manpower too. As it is automatized, it can maintain the accurate time, which follows perfect rules, exact bill and many more flexibilities. In this article, we discuss artificial intelligence, and we propose an idea on how to use this artificial intelligence to avoid the abovementioned malpractices and problems.

4.1 Introduction

In this busy world, everyone is bothering about their time, nowadays, time is very precious, and people are forgetting about their regular bill payments (like electricity). Presently, the electricity billing is manual and time-consuming process. With the proposed cloud-based billing system, we can overcome the disadvantages which are

P. Pramod Kumar (✉)
Department of Computer Science and Artificial Intelligence, SR University, Warangal, Telangana, India

N. Achha
Department of Electronics and Communication Engineering, S R Engineering College, Warangal, Telangana, India

G. Ranadheer Reddy · V. Pranathi
Department of Computer Science and Engineering, Sumathi Reddy Institute of Technology for Women, Warangal, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_4

occurring due to the present manual billing system [1]. In this cloud-based system, there is no need of producing the bill and distributing the bill manually [2].

The main idea of this paper is to give the best and well-automated electricity billing and payment management system; this makes our lives easy in this faster generation. This system reduces or eliminates the paper wastage and manpower; to pay our bills using this system, we are required to have a fourth-generation (4G) subscriber identity module (SIM) [3]; this 4G SIM will be linked with the Aadhar number of the person on whose name the electric meter number was issued. With the help of the Aadhar number, details of the person will be accessed. The SIM which is fixed to the meter reads the number of units consumed by the consumer, and then it automatically calculates the charges. The consumer can easily pay the bill amount online through various e-payment gateways and Web sites. With this kind of system, no one can escape from the fines or punishments. First of every month consumers will automatically get the amount to be paid through the SMS, e-mail and departmental applications which are provided. This service perfectly follows the government rules like charges per unit, taxes and punishments (like disconnecting the power supply if consumer fails to pay after 15 days of the bill issued). The consumer will be allowed to pay the amount more than the bill generated, but they cannot pay the less amount than the bill generated, and extra paid amount will be added to the adjustment charges to reduce from the next month bill.

4.2 Related Work

Many people have worked on similar kind of systems in different ways and proposed versatile methodologies to avoid the problems faced in the manual billing process. In the present traditional system of bill generation, one employee from electricity department visits to each and every house in that area with a billing machine provided by the department [4]. He places that billing machine in front of the meter, then it automatically scans, counts the number units consumed and prints the automatically calculated bill including different taxes and charges, and finally, a receipt of paper will be issued to the consumer. However, the bill is calculated automatically but here still the manpower is required and paper wastage is done. To overcome this type of problems, we are making automatized billing, where the manpower is not required and paper wastage is reduced. However, raw details stored in the cloud storage and bill can be printed as on the request [5]. The below given figure is one such kind of smart billing system published in July 2015 (Fig. 4.1) [6].

The below figure is one such kind of a proficient and smart electricity billing published on July 13, 2019 (Fig. 4.2).

Here, a work is done that displays the warning to pay the bill on the light-emitting diodes (LED) fixed to it. The block diagram of prepaid energy meter is shown in Fig. (i). It consists of microcontroller LPC2148, buzzer, relay, single-phase energy meter, MAX232, infrared (IR) sensor circuit and liquid crystal display (LCD) [2] (Figs. 4.2 and 4.3).

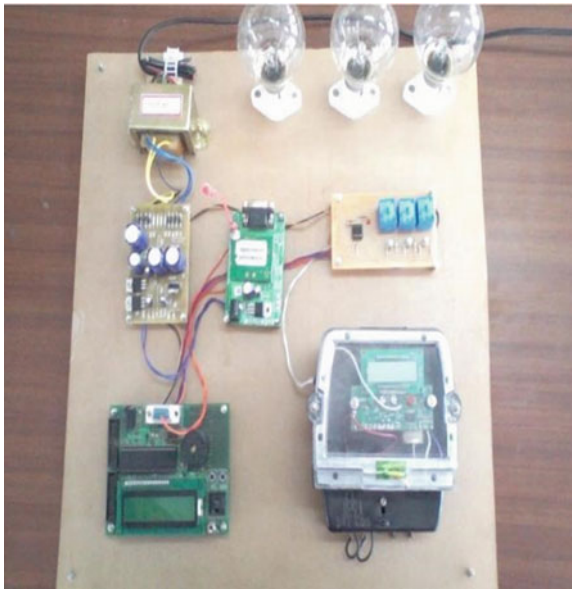


Fig. 4.1 Smart billing system published in July 2015

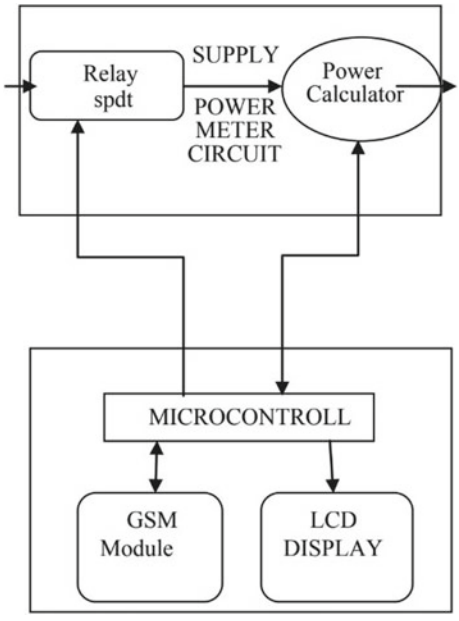


Fig. 4.2 Proficient and smart electricity billing published on July 13, 2019

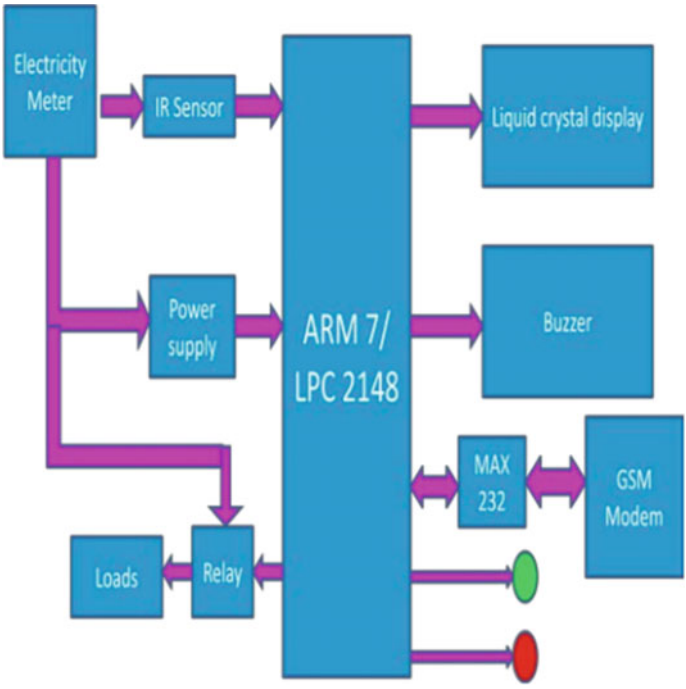


Fig. 4.3 Smart electricity billing system from volume-6, issue-3, May–June 2016

4.3 Methodology

We discuss the proposed methodology in three different sections, as follows.

4.3.1 Automatic or Calculation Section

Electric meter will be fixed with a 4G SIM [3, 7], which is linked with the Aadhar number of a consumer on whose name the meter is registered. If the house owner wants to change the details or wants to update any thing, they have to show the perfect proofs as documents, so that details of consumer will be updated or changed [8, 9]. In the beginning of the every month, the details of the consumer including the meter reading will be sent automatically to the cloud space with the help of the SIM and the data will be stored in the cloud [10, 11, 12]. The meter number, meter reading and other required details will be sent to the automatic calculator, and then the calculator will generate the bill which includes all taxes and charges. Then, the final bill amount will be sent back to the cloud storage (Fig. 4.4).

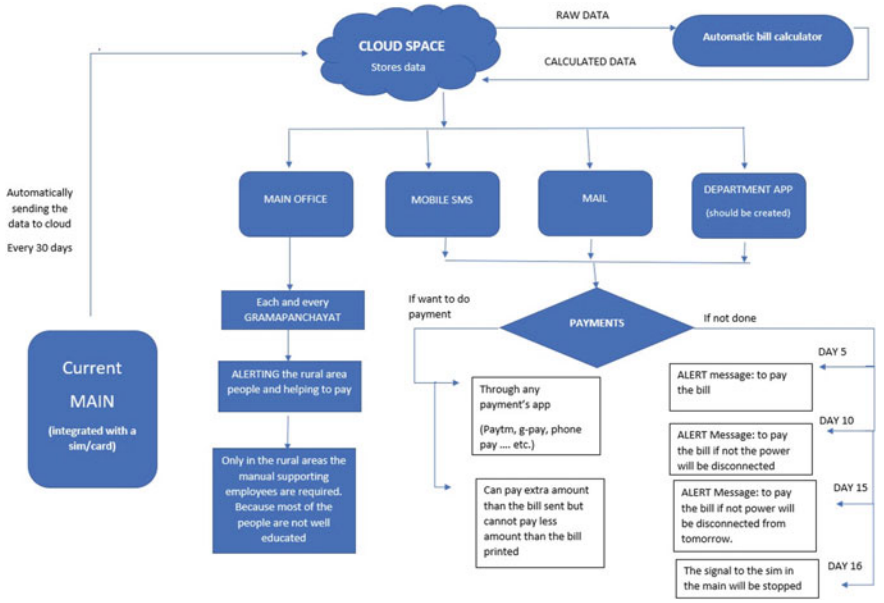


Fig. 4.4 Cloud-based centralized smart electricity billing and payment management system

4.3.2 Payment Section

From the cloud storage, bill data will be sent to the electricity department head office database and to the consumer through the mobile SMS and e-mail which are linked to Aadhar number with a payment link added to it [10, 11, 13]. The app published by the electricity department will be notifying the consumers with payment details. When the consumer wants to pay, then the consumer will be directed to the online payment Web site. If consumer fails to pay the bill, then alert messages will be sent, the first alert message will be on 5th day after the bill issuing date to remind to pay the bill before the last date mentioned, the second alert message will be sent on 10th day after the bill issuing date to remind to pay the bill and to warn about the power supply disconnections, and the third alert message will be sent on the 15th day after the bill issuing date, that is to remind bill payment. After three alert messages, if no payment is made by the consumer, then by the next day automatically the electric supply will be disconnected by disabling the SIM inside the meter. If the consumer wants to regain the power connection, they must have to pay the previous pending bills and have to put a request for connection resume [6, 14, 15].

During the payment of bill, the consumer is directed to the secured payment Web sites, and there the consumer will have an option to edit the amount to be paid (adjustment charges) [16, 17]. They can pay the amount more than the bill generated, and that extra paid amount will be adjusted for the future consumptions. But, they cannot pay the amount less than the bill amount generated. If the payment

is done before due date, the connection is stable; otherwise, the connection will be disconnected by deactivating the SIM automatically.

This is a fully automatic process where no one can escape from the punishments and the late payments. Since this is a fully automatic system with all functions, the urban area consumers can do their payments through the e-payment services as they are aware. From the electricity department side, if any charges are changed (like charge per unit, taxes, etc.), each and every consumer will be intimated through the SMS and e-mails about the new regulations.

4.3.3 Rural Areas Billing

When it comes to the scenario of rural areas, all the consumers are not mostly literates and everyone may not be aware of using smart phones and e-payment applications. So, the manual process should be done by local authorities to help the illiterate and old people on the first of every month. As the details from the cloud storage are sent to the electricity central office, from there the village wise consumer list along with generated bill amount for that current month will be sent to respective local electricity offices and that list will be made available on notice board. Hence, everyone can see that list, and by taking the office employee help, rural consumers also can pay bills in time. Those employees should take care about the alerts that to be given manually by moving to each and every house, whose bills are not paid on every 5th, 10th and 15th day after the bill issued [6, 18, 19].

However, the most of the work is done automatically with the help of the SIM and cloud storage. It can be used in the rural areas also. By this we can reduce the manpower utility and usage of the paper.

4.4 Performance Analysis

4.4.1 Comparison of Smart Meter with Traditional Meters

The below table gives the differences between smart meter and traditional meters, and as per the properties listed we can say that the performance of smart meter is much more better than traditional and helpful to both electricity department and consumers (Table 4.1).

Table 4.1 Comparison of smart meter with traditional meters

Property	Electromechanical meter	Digital meter	Smart meter
Accuracy	Low	High	Very high
Theft detection	Low	Possible at node level	Possible at network level
Communication	Manual reading	One-way communication	Two-way communication
Control	No	Limited	Full
Consumer participation	Nil	Less	High
Time wise reading	No	After a fix interval	According to requirement
Day-to-day billing info	Not possible	Not possible	Possible

4.4.2 Research Limitations and Practical Implications

If this automatic system is brought to existence, there will be a lot of changes in our activities. It reduces the paper wastage and manpower utility. Many miscalculations can be avoided, and strict punishments will be executed. However, in rural areas more than the 50% of the work is done automatically. As this is useful to electricity department and to the consumers, it can be implemented as soon as possible. There will be a helpline services for 24*7 to rectify the problems and to provide help [20].

4.4.3 Originality and Outcomes

In this way, we can use the artificial intelligence to decrease the manpower utility and to decrease the malpractices. After using this artificial mechanism, these are some outputs we get.

- We can do these things with less a cost.
- It decreases the usage of the paper.
- This mechanism can be used in anywhere irrespective of the area (village, town/city)
- This does not require the heavy manpower.
- Malpractices cannot be done.
- Strict rules will be followed by the program where no one can miss the rules regarding punishments or anything.
- Changes in the charges can be done directly through the server.
- Changed charges and details will be informed and notified through the SMS and e-mail services.

- Through SMS, e-mails and app notifications, the bill and payment details will be intimated.
- While paying the amount, we can pay the amount more than the bill issued but cannot pay the bill less than the bill issued, and those extra money will be adjusted for future balance.
- There will be the app which is more helpful.
- It is developed with a user-friendly environment.

4.5 Conclusion and Future Scope

In this well-turned out society, we need smart techniques to get rid of the problem. This kind of technology of artificial intelligence is the solution for many problems such as, wastage of manpower, incompetent and inaccurate billing, profusely increasing dereliction of duty and asymmetrical payments in electricity billing departments. Proper execution of tax and bills will amplify nation's economy.

In this article, we have given a solution to the problems which we have mentioned in the abstract. If we use this type of artificial intelligent mechanism, we can make our day-to-day life easy. The human power which we are using now can be reduced. We can make better understanding of the power consumption. We can provide the accurate billing and those customers who have these meters installed and want to have strict control over monthly bills can also go for prepayment using smart billing tools. From the side of government, they can have the control over the revenue, since this also works in the rural areas.

In future, if government wants to the setup to be easy and user friendly, they should develop an app or update the version of present available mobile app to make easy access to online payments and in that app we can also be able to file complaints, apply for new meter connections, etc.

Funding No funding sources.

Conflict of Interest Authors have declared that no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Senavirathna, A.M.N.N., Wijesinghe, W.A.S., Premachandra, C.: Smart home energy management system to reduce monthly electricity bill. In: 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), pp. 243–245. Osaka, Japan (2019). https://doi.org/10.1109/GCCE46687.2019.9015221
2. The Bible of electric meters, continuously updated since electricity was discovered. In: Handbook for Electricity Metering. The Edison Electric Institute

3. Jain, A., Kumar, D., Kedia, J.: Smart and intelligent GSM based automatic meter reading system. Int. J. Eng. Res. Technol. (IJERT) **2**(3),1–6. ISSN: 2278–0181
4. Leung, K.K., Yung, N.H.C., Cheung, P.Y.S.: Novel neighborhood search for multiprocessor scheduling with pipelining. In: Proceedings Fourth International Conference/Exhibition on High Performance Computing in the Asia-Pacific Region, vol. 1, pp. 296–301. Beijing, China (2000). https://doi.org/10.1109/HPC.2000.846565
5. Kanthimathi, K., et al.: GSM based automatic electricity billing system. Int. J. Adv. Res. Trends Eng. Technol. (IJARTET) **2**(7), 16–21 (2015)
6. Through SMS, e-mails and app notifications the bill and payment details will be intimated.
7. Pramod Kumar, P., Sagar, K.: A proficient and smart electricity billing management system. In: 1st International Conference on Emerging Trends in Engineering (ICETE 2019), held in Hyderabad, pp. 156–160 (2019)
8. Kumar, K.S., Bhavana, J.: A study on data mining towards cloud computing. Indian J. Public Health Res. Develop. **9**(11) (2018)
9. Pramod Kumar, P., Thirupathi, V., Monica, D.: Enhancements in mobility management forfuture wireless networks. Int. J. Adv. Res. Comput. Commun. Eng. **2**(2) (2013)
10. Kundeti, K., Pallagani, S.: Smart electricity billing system. Int. J. Eng. Manage. Res. 526–529
11. Alam, M.J., Shahriar, F.M.: Electricity billing systems at residential level in sylhet city: is pre-paid system perceived as a better option by the subscribers? Indust. Eng. Lett. **2**(3) (2012). ISSN 2224–6096 (print) ISSN 2225–0581 (online)
12. P. Pramod Kumar, S. Naresh Kumar, V. Thirupathi, Ch. Sandeep, "QOS AND SECURITY PROBLEMS IN 4G NETWORKS AND QOS MECHANISMS OFFERED BY 4G", International Journal of Advanced Science and Technology, Vol. 28, No. 20, (2019), pp. 600–606.
13. Rastogi, S., Sharma, M., Varshney, P.: Internet of Things based smart electricity meters. Int. J. Comput. Appl. (0975–8887) **133**(8):13–164 (2016)
14. Rathnayaka, M.R.M.S.B., Jayasinghe, I.D.S., et al.: Mobile based electricity billing system (MoBEBIS). Int. J. Sci. Res. Publ. **3**(4):1–5 (2013). ISSN: 2250–3153
15. Sandeep, C.H., Naresh Kumar, S., Pramod Kumar, P.: Security challenges and issues of the IoT system. Indian J. Public Health Res. Develop. **11**, 748–753 (2018)
16. Shinde, M.V., Kulkarni, P.W.: Automation of electricity billing process. In: 2014 International Conference on Power, Automation and Communication (INPAC), pp. 18–22. Amravati (2014). https://doi.org/10.1109/INPAC.2014.6981128
17. Shoults, R.R., Chen, M.S., Domijan, A.: The energy systems research center electric power system simulation laboratory and energy management system control center. IEEE Power Eng. Rev. PER 7(2), 49–50
18. Tamarkin, T.D.: Automatic Meter Reading. Public. Power Mag. **50**(5), 934–937 (1992)
19. Tariq J.: Design and implementation of a wireless automatic meter reading system. In: Proceedings of the World Congress on Engineering (2008)
20. Mohale, V.P., Hingmire, A.G., Babar, D.G.: Ingenious energy monitoring, control, and management of electrical supply. In: 2015 international conference on energy systems and applications, pp. 254–257. Pune (2015)

Chapter 5
Real-Time Driver Distraction Detection Using OpenCV and Machine Learning Algorithms



V. N. V. L. S. Swathi, D. Akhilesh, G. Senthil Kumar, and A. Vani Vathsala

Abstract Continuous driver interruption identification is the center to numerous interruption countermeasures and major for building a driver-focused driver help framework. As more remote correspondence, diversion and driver help frameworks multiply the vehicle advertise, and the rate of interruption-related accidents is relied upon height. This article presents a diagram promising methodology which is to grow continuous driver interruption countermeasures, including three categories: distraction prevention before distraction occurs; distraction mitigation after distraction occurs; collision system adjustment when a potential collision is estimated.

5.1 Introduction

Nowadays, many accidents occurring due to driver distraction play important role. According to National Highway Traffic Safety Administration (NHTSA) and World Health Organization (WHO), consistently more than 1.24 million individuals kick the bucket in car crashes far and wide, which are among children, heavy vehicle driver, due to proper sleep. In this venture, we plan to build up a driver interruption discovery framework which is utilized to distinguish the driver from threat. In this system, we are going detect whether the driver is feeling sleepy or not.

In the proposed system, the vehicle is tracked continuously and monitored using the system which contains Raspberry Pi 3, pi camera. If the driver distraction is due to sleepy, wrong path, the alert system will active, then status of the vehicle is sent to the user as alert message via mail (or) message.

The paper is organized as follows: Section 1 contains the introduction, Sect. 2 contains the background, Sect. 3 contains the related study, Sect. 4 contains the

V. N. V. L. S. Swathi (✉) · D. Akhilesh · A. V. Vathsala
CVR College of Engineering, Hyderabad, India

G. Senthil Kumar
SRM Institute of Technology, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_5

methodology, Sect. 5 contains the Implementation, Sect. 6 explains the results, and Sect. 7 describes conclusion and future work.

5.2 Background

5.2.1 Real-Time System [1]

Real time is the very short amount needed for computer system to receive data and information and then communicate it or make it available.

The continuous working framework utilized for a constant application implies for those application where information preparing ought to be done in the fixed and little quantum of time. It is not quite the same as broadly useful PC where time idea isn't considered as much critical as progressively working framework.

5.2.2 Driver Distraction Detection System [2]

Mainly driver interruption frequently brought about by following elements: sleep, work period, work pressure and actual individuals regularly attempt to accomplish work much in a day and they lose valuable rest because of this. They take coffee, tea or different simulant individuals keep on remaining alert that happens the body at long last implodes and the individual falls rest. The variables caused the driver interruption, and human mind is prepared to think there are times the body should be sleeping; basically this happens during the hour of sunrise and sunset. Between the long periods of 1–6 am, this reason truly ill suited, by being either under or overweight will cause weakness. In this location, framework was clarified in methodology.

5.2.3 OpenCV

OpenCV (open-source computer vision library) supports a lot of computer vision and machine learning application. Mainly, OpenCV supports programming languages like Python, C++ , Java. In Python, OpenCV Python API and other libraries like SciPy, Matplotlib and NumPy can be used with this.

5.3 Related Study

Related study is also known as literature survey; it is most important step in development of the software process. Before developing the software, it is necessary to determine which platform can be developed for our project (driver distraction detection system). In the previous long term, numerous explorers have been chipping away at the advancement of driver distraction detection system to forestall street mishap by utilizing various strategies. The ongoing methods depend on physiological measures like brain waves, pulse, heart rate… … and so on,

We have referred many papers; from those papers we understood what are the limitation of the existing system. One of the existing system [3]. According to the experts, it has been observed that when the driver does not take break, they tend to run a high risk of becoming drowsy; this study shows that accidents occur due to sleepy driver in need of a rest, which implies that street mishaps happen more because of sluggishness as opposed to drink driving. Consideration help can warm of in mindfulness and laziness their present status of exhaustion and the driving time since in the last break.

The most reinforcement and recuperation procedures that have been created in driving discovery strategies, for example,

ECG and EEG [4]
Electrocardiogram (ECG), electroencephalogram (EEG) are the physiological sign most usually to gauge tiredness. The ECG and EEG signals have different frequency bands and range. Its disadvantages are implementation and complexity.

Steering Wheel Movement (SWM) [5]
It is estimated utilizing directing point sensor, and it is broadly utilized for vehicle-based measure for the recognizing the degree of driver tiredness. To eliminate the effects of lane changes, the researches thought about which are expected to change the horizontal situation inside the path. Model a vehicle organizations, for example, Nissan and Renault have received SWMs yet it works in exceptionally restricted circumstances. Its disadvantages are costly technique, redundancy issue, and complexity.

5.4 Methodology

In this work, we have characterized the Driver Distraction Detection Algorithm. We utilize the camera-based framework to notice his visual conduct and the vehicle sensor to quantify the driving conduct. According to development of real-time driver distraction countermeasure includes three categories by using it.

5.4.1 Distraction Prevention Before Distraction Occurs

In this method, we are using to prevent distraction before distraction occurs, the vehicle real-time identification measure by recording the video arrangement of the drivers. The video transfers got from web camera are driven the framework, starting doesn't known the underlying situation of the face in light of the fact that the eye squint is a fast closing and continuing of a characteristic eye. Each individual has fairly one of a kind case of squint, the model shifts in the speed of closing and opening, a degree of pounding the eye and in a glimmer range the eye squint props up approximately 100–400 ms.

The framework gets the main picture and attempts to discover the face district in the picture utilizing the histogram of oriented gradients and linear SVM object indicator algorithm explicitly for the errand of face location.

To build our face detection, the author used Viola–Jones algorithm. By comparing both algorithms, the HOG [6] accuracy is more. It is used to prepare higher-precise item classifier in their specific investigation, human identifiers.

HOG

A HOG is a component descriptor by and largely utilized for object location. A HOG depends on the property of items inside in a picture to have the conveyance of power angles or edge headings. Slopes are determined inside a picture for each squares. A square is considered as a pixel lattice wherein angles are established from the size and heading of progress in the forces of the pixel with in the square.

HOG figures the classifier for each picture in and size of the picture, applied the sliding windows, removed HOG descriptor at every window and applied classifier. All the face test photos of an individual are dealt with to the part descriptor extraction figuring, i.e., a HOG. The descriptors are angle vectors produced per pixel of the picture. The slope for every pixel comprises of size and heading, determined utilizing the accompanying formulae:

$$\begin{aligned} \mathrm{g} &= \sqrt{g_x^2 + g_y^2} \\ \theta &= \arctan\frac{g_y}{g_x}, \end{aligned} \tag{5.1}$$

In the current formulae, G_x and G_y are separately the level and vertical segments of the adjustment in the pixel power. A window size of 96×144 is utilized for face pictures since it coordinates the overall viewpoint proportion of human countenances. The descriptors are resolved over squares of pixels with 8×8 estimations. These descriptor esteems for every pixel more than 8×8 square are quantized into 9 canisters, where each receptacle speaks to a directional point of slope and incentive in that container, which is the summation of the extents of all pixels with a similar point. Further, the histogram is then standardized over a 16×16 square size, which implies four squares of 8×8 are standardized together to limit light conditions. This instrument mitigates the precision drop because of an adjustment in light. On

the off chance that the classifier identified an item with enough likelihood that takes after a face, the classifier recording the bouncing box of the window and applied non-greatest concealment to make the precision expanded (Fig. 5.1).

Linear SVM [7]
Support vector machines (SVMs) or backing vector machines (BVMs) are managed AI models that partition and order information. The SVM model is readied using different HOG vectors for various appearances (Fig. 5.2).

Facial Landmarks Detection with Dlib
The Dlib C++ library is a cross-stage bundle for stringing, organizing, numerical activities, AI, PC vision, and pressure, setting a solid accentuation on very and compact code. From a PC vision point of view, Dlib has various best in class executions, including facial milestone discovery, correlation following, deep measurement learning. Facial landmarks they are used in computer vision applications. Recognizing facial places of interest is a subset of the shape conjecture issue. Given an info

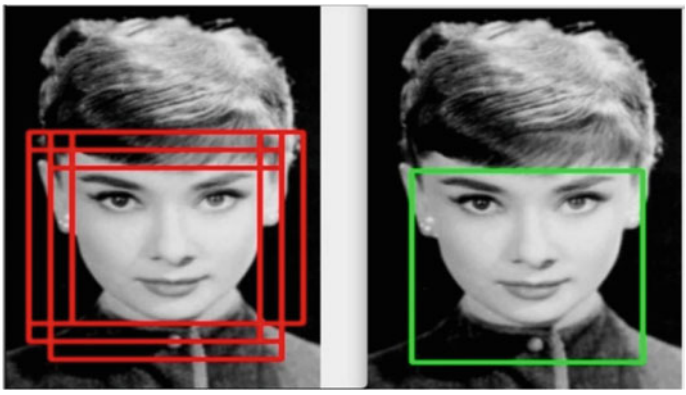


Fig. 5.1 (Left) Detecting numerous covering bouncing boxes around the face we need to identify. (Right) Applying non-greatest concealment to eliminate the excess bouncing boxes

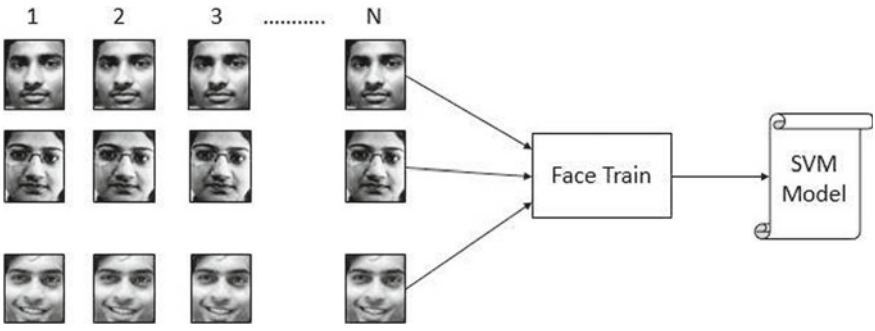


Fig. 5.2 Face training on linear SVM

picture for (HOG + VSVM) and regularly a ROI that determines the object of interest a shape predicator endeavors to neighborhood key focal points along the shape. With respect to facial achievements, our goal is recognize critical facial structures on the face using shape desire methods.

Recognizing facial tourist spots is subsequently a two-stage measure:

Stage 1: Point the part of the face in the image.
Stage 2: Finding key facial structures on the face region. We may apply a pre-arranged HOG + Linear SVM object identifier unequivocally for the endeavor of face acknowledgment. In one or the other case, the certifiable figuring used to perceive the face in the image doesn't have any kind of effect. Or maybe, what's huge is that through some methodology we get the face ricocheting box (i.e., the (*x*, *y*)—bearings of the face in the image). Given the face locale we would then be able to apply recognizing key facial structures in the face area. There are a grouping of facial achievement pointers, yet all strategies fundamentally endeavor to limit and name the going with facial locale: right eye, left eye, nose, mouth.

The facial landmark detector and Dlib library of the paper [8].

This strategy begins by utilizing: A preparation set of named facial tourist spots on a picture. These pictures are physically marked, indicating explicit (*x*, *y*)—directions of locales encompassing facial structure. Priors, of even more expressly, the probability on partition between sets of information pixels. Given this readiness data, a social occasion of backslide trees is set up to survey the facial achievement positions genuinely from the pixel powers themselves (i.e., no "feature extraction" is happening). The end result is a facial achievement identifier that can be used to recognize facial places of interest dynamically with first rate conjectures.

Advance Dlib's Facial Milestone Indicator

The advance facial milestone indicator and Dlib library are used for 68 (*x*, *y*)—coordinates on the face. The below image describes the 68 coordinates of Dlib's facial landmark detector (Fig. 5.3).

Eye Blink Detection

To fabricate our blink identifier, we'll be registering a measurement called the eye aspect ratio (EAR), presented in paper [9].

Not in any way like traditional picture planning systems for enrolling squints which regularly incorporate a mix of eye repression. Threshold value tells about the whites region of the eyes. Deciding whether the "white" region of the eyes vanishes it showing a squint. This procedure for eye glint acknowledgment is speedy, capable, and easy to execute.

As far as flicker identification, we are simply enthused about two plans of facial structures the eyes. Each eye point is spoken by 6 coordinates, in vertical and horizontal axis. From the point p1 of the eye, we make point of the location which are working clockwise region around the eye.

In the above Fig. 5.4, we should eliminate on key point. There is a connection between the width and the stature of these directions. In view of the work [5], we



Fig. 5.3 68 coordinates of Dlib's facial landmark

would then be able to infer a condition (2) that mirrors this connection called the eye aspect ratio (EAR):

$$\text{EAR} = \frac{||p_2 - p_6|| + ||p_3 - p_5||}{2||p_1 - p_4||}, \tag{5.2}$$

In the above Eq. (5.2) $p_1, p_2, p_3, p_4, p_5, p_6$ are 2D facial milestone areas. The numerator of this condition calculates the detachment between the vertical eye places of interest while the denominator enrolls the partition between level eye achievements,

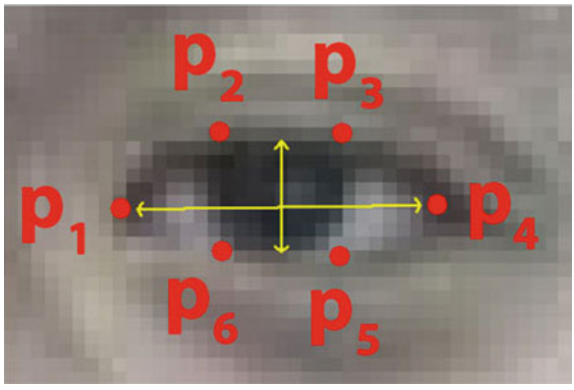


Fig. 5.4 6-point location of the eye

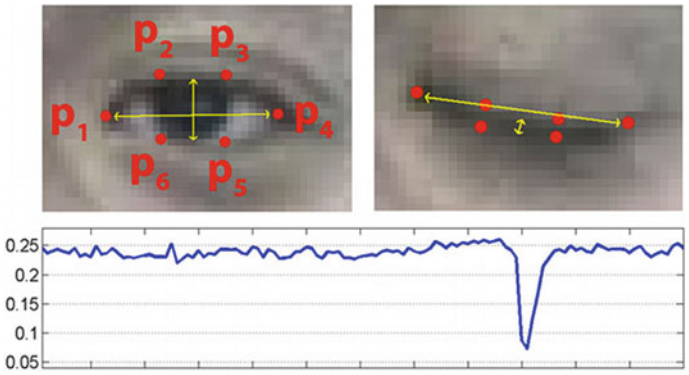


Fig. 5.5 6-point location of the eye is open. Eye indicator when the eye is closed. Graph of the eye aspect ratio over time

weighting the denominator fittingly since there is only one parcel of even concentrates yet two plans of vertical core interests. For sure, as we'll find, the eye viewpoint extent is around consistent while the eye is open, anyway will rapidly tumble to zero when a squint is happening. Utilizing this basic condition, we can dodge picture preparing procedures and just depend on the proportion of eye milestone separations to decide whether an individual is flickering.

To make this more clear, consider the following Fig. 5.5.

In Fig. 5.5 top-left, we have an eye that is totally open the eye perspective extent here would be large and decently reliable after some time. Be that as it may, when the individual squints (upper right) the eye perspective proportion diminishes drastically, moving toward zero. The down graph tells about the eye angle proportion after some time for a video cut. As should be obvious, the eye viewpoint proportion is steady, at that point quickly drops near zero, at that point increments once more, showing a solitary flicker has occurred.

In our next segment, we'll figure out how to execute the eye angle proportion for flicker location utilizing facial tourist spots, OpenCV, Python, and Dlib.

Drowiness Detection Results

To run the above detection, we are writing a program code on Python, to run this program on your own Raspberry Pi, as shown in system architecture Fig. 5.6. The detection system checks for driver drowsy or not. If drive is drowsy an alarm is generated to alert the driver who is in a drowsy state, otherwise the process is terminated.

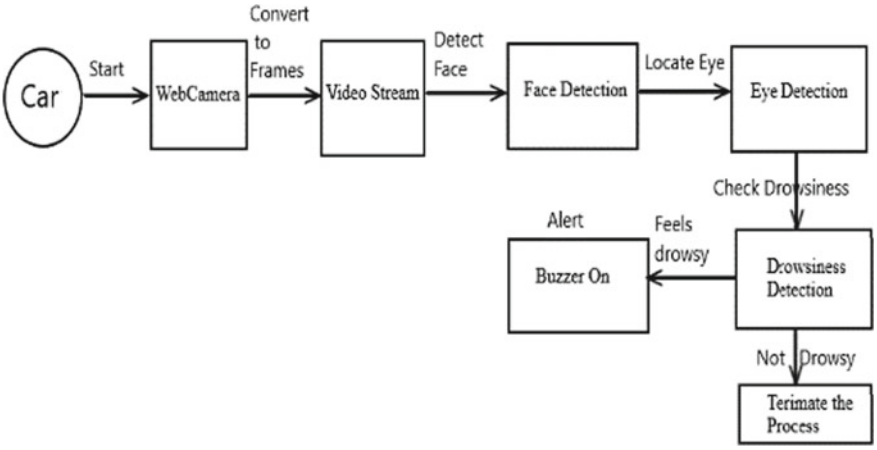


Fig. 5.6 System architecture

5.4.2 Distraction Mitigation After Distraction Occurs

In this case were going implement the driver alert system through send message and location of the car, in the case of continuous drowsy of a person.

5.4.3 Collision System Adjustment When a Potential Collision

The system can be enchased with more additional sensors to improve the safety of the users in the vehicle like speed limit, 360 degree camera and automatic sensors.

5.5 Implementation

The tools are connected in this way, so that we can collect the data. The tools used are Raspberry Pi 3, pi-camera, GPS, and USB (Fig. 5.7).

The process of collecting the data from the system that contains Raspberry Pi 3, pi-camera, GPS is done in the following way:

1. The webcam is used mainly to capture the static image or the video stream. This captured image or video stream is given as contribution to the framework.
2. Captured image processing module will process the obtained image or video input fed into the system. It processes the input and converts it into frames for further processing.

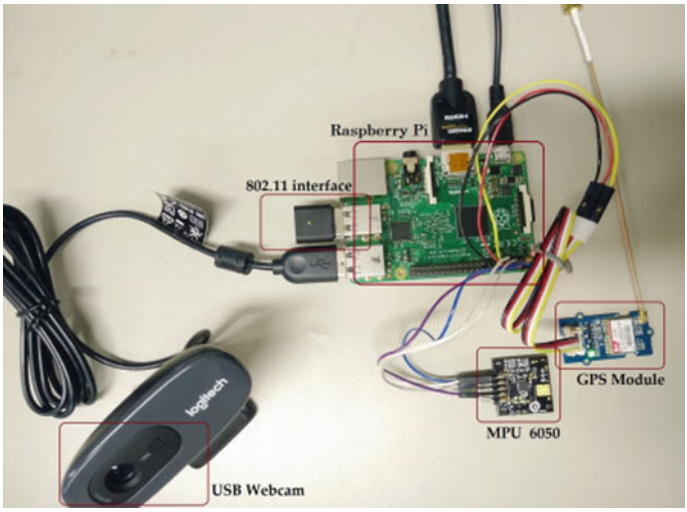


Fig. 5.7 Proposed system architecture kit

3. The face detection technique is used to locate the face from the image. The video stream which is given as input is converted into frames, and the face is detected from those frames. Haar classifier and HOG + SVM algorithm is used to detect the face from a given image.
4. The situation of the driver's eye is controlled by utilizing suitable limit. In this work, edge area of the eyes region is considered.
5. After locating the eye region in the frame, the system finds out whether the eye is in a closed state or in open state. In the event that the eye stays open, at that point the framework gives the message that the driver isn't feeling languid. On the off chance that the eyes stay shut, at that point the framework gives the message that the driver is feeling sluggish.
6. The output from the system includes: Alert Message to the user via buzzer and terimate the process. The indication is given in order to prevent road accidents which occur due to the drowsiness of the driver.

5.6 Result

Our framework created the ready when the driver gets weakness and on the caution, gets ready message, area of vehicle to the client precisely when driver gets occupied.

The system is tested, and the results obtained are furnished below:

1. Checking for the driver distraction as shown in Fig. 5.8.
2. Giving alert when driver gets distracted as given below Fig. 5.9.

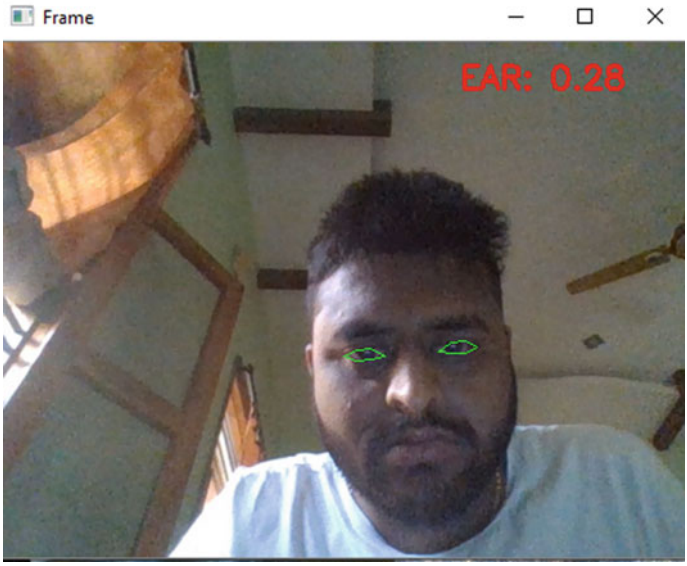


Fig. 5.8 Checking for drowsiness

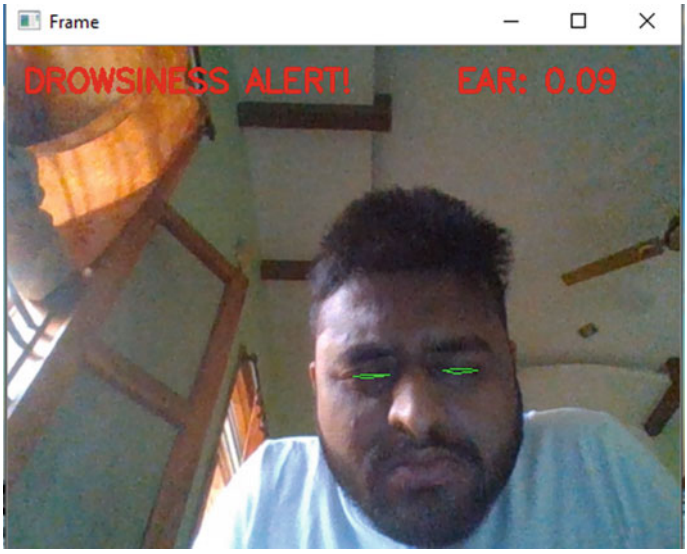


Fig. 5.9 Alert

5.7 Conclusion and Future Work

In this paper, driver distraction detection system will efficiently find the drive fatigue and sends alert message, location of vehicle to the user. The system can be enchased with more additional sensors to improve the safety of the users in the vehicle. The invoice generation can also be implemented as additional feature in future.

References

1. Zhang, H., Smith, M.R.H., Witt, G.J.: Identification of real-time diagnostic measures of visual distraction with an automatic eye-tracking system. Human Factors. **48**(4), 805–821 (2006)
2. Liu, T., Yang, Y., Huang, G.-B., Yeo, Y.K., Lin, Z.: Driver Distraction Detection Using Semi-Supervised Machine Learning. IEEE Trans. Intell. Transp. Syst. **17**(4), 1108–1120 (2016). https://doi.org/10.1109/tits.2015.2496157
3. VandnaSaini, R.S.: Driver Drowsiness Detection System and Techniques: A Review. Int. J. Comput. Sci. Informat. Technol. **5**(3), 4245–4424 (2014)
4. Singh, K., Kaur, R.: Physical and physiological drowsiness detection methods. IJIEASR, **2**, 35–43 (2013)
5. Krajewski, J., Sommer, D.: Steering wheel behavior based estimation of fatigue. In: Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Germany
6. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 886–883. San Diego, CA, USA (2005). https://doi.org/10.1109/CVPR.2005.177
7. Tomasz, M., Gupta, A., Efros, A.A.: Ensemble of exemplar-SVMs for object detection and beyond. In: 2011 International Conference on Computer Vision, 89–96 (2011)
8. Kazemi, V., Sullivan, J.: One millisecond face alignment with an ensemble of regression trees. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1867–1874. Columbus, OH (2014). https://doi.org/10.1109/CVPR.2014.241
9. Tereza, S., Cech, J.: Real-Time Eye Blink Detection using Facial Landmarks (2016)

Chapter 6
Data-Driven Prediction Model for Crime Patterns



Y. Gayathri, Y. Sri Lalitha, M. V. Aditya Nag, and Sk. Althaf Hussain Basha

Abstract With the ever-increasing technology widgets and media, there is enough awareness of good and bad in the surroundings. Unemployment or abundant money is making people to indulge in crime. Throughout the world, crime prediction and prevention are the two major issues of all governments for the safety and security of their citizens. With machine learning and data mining techniques to predict unknown events, crime data analysis and prediction has attracted the attention of many researchers in recent times. There is still generalized, accurate and optimized requirement of crime data analysis and prediction models. This work proposes to address crime data analysis and prediction using visual analytics and machine learning algorithms on Kaggle crime dataset. The proposed models provide better insights on the crime zones and frequency of occurrences of certain crimes in a zone. We have compared Naïve Bayesian and Random Forest Machine Learning Models to predict the crime, and we have noticed that our models exhibited better accuracy.

6.1 Introduction

Crime is an un-lawful act punishable by state or authority. Crime is an act harmful not only to some individual but also to a community, society or the state. Therefore it is socio-economic problem. The nature of the crime varies depending on society, education, unemployment and poverty levels. The types of crimes such as assault, aggravated assault, burglary, murder, rape, kidnapping and stabbing are common and occur periodically. Knowing in advance the occurrence of crime and preventing it is a great help to society and man-kind. Predicting a crime type or its reoccurrence

Y. Gayathri (✉) · Y. Sri Lalitha
Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

M. V. Aditya Nag
Institute of Aeronautical Engineering, Dundigal, Hyderabad, India

Sk. Althaf Hussain Basha
A1, Global Institute of Engineering and Technology, Markapur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_6

is not trivial. The advancements in machine learning techniques and open-source tools to implement the same is making it possible to provide solutions to many real-time problems. There is a vibrant research community contributing improved models in crime data analysis and prediction. The generalized, accurate models are yet to be generated. The accurate prediction models use data and statistical analysis for predicting crime types and are useful for law-enforcement departments.

In this work, we explored visual analytics for crime data analysis and machine learning models to predict crime categories. In the process, we used R language an open-source statistical tool to build visualization and prediction models. With the high volumes of data storage mechanism, we can record all the crimes and the criminal activities in structured and un-structured forms. To derive the required knowledge from this voluminous data, visualization tools are most helpful. It is impossible to know the characteristics of data without visual analytics. Visualization helps in performing exploratory data analysis and identifies required features for building a prediction model. R language provides extensive visualization functionalities; we explored this here to determine high-crime-prone locations, crime types, days of a week of high crime activity and the time of its occurrence. The machine learning algorithms such as Naïve Bayesian and Random Forest Prediction Models are implemented for predicting a crime categories.

6.2 Literature Survey

Literature study on crime data analysis and crime prediction significant work has taken place. The researchers proposed solutions ranging from application of statistics, visual analytics, machine learning, data mining, GIS-based analysis to artificial neural networks [13]. Still this area requires *generalized, accurate and optimized solutions constituting* various aspects of crime data recorded. This section briefs certain previous works proposed.

In many research papers, statistics is applied to analyze crime data for frequent crime types, crime zones and applied visual analytics to get insights of crime data [12]. In [1], human social data obtained from social media via mobile activity combined with demographic information using real crime data to predict crime hotspots in London, UK. Machine Learning methods such as Naïve Bayesian, Decision Tree using WEKA an open-source data mining technique applied in [2]. Considering characteristics such as weather, road, driver and car conditions, the occurrences of road accidents were explored in [3] using K-NN, Naïve Bayesian and decision trees algorithms of data mining with 79–81% accuracy. At University of Arizona, the data mining techniques were applied to identify the crime zones [5, 8, 9]. Reviews different crime-prediction techniques and suggested that statistics, ML, AI and Data Mining are the effective mechanisms in crime prediction. In [10], Gradient Boost and Page Rank techniques were applied to determine and destroy the hidden links in criminal networks. GIS-based analytics is applied in [11] to identify areas of crime and alert

cops for necessary actions. A comparative study of linear regression, additive regression and Decision stump algorithms using WEKA is made for crime prediction in [4]. In [7] drug-related crime data and emerging crime zones were identified using visual and spatial analytics. Mobile network activity and demographic information together used to predict crime zones in UK. This work uses visual analytics with R to analyze the crime data and predict crime type using naïve Bayes and random forest approaches.

6.3 Our Approach

This work involves the steps like dataset collection, preprocessing, analyzing nature of dataset using visualization for feature selection [6], building a prediction model and evaluating the models with unknown criminal event.

Our process of building a prediction model is depicted in Fig. 6.1.

Dataset Collection
We have considered Kaggle Crime dataset which constitutes data from 2001 till date. It is a collection of criminal and traffic offenses in the country. The work here considered data from 2012 to 2017. It constitutes around 457,000 crime records. From this collection after preprocessing, we have considered 333,830 records of different offense categories for experimenting.

Data Preprocessing
The dataset is composed of 19 attributes. Not all of these attributes are relevant for our analysis. Only the relevant attributes are considered for efficient performance. The dataset is cleaned by removing null valued rows and invalid dates. Selected records randomly for each offense_category thereby forming a preprocessed dataset. The dataset consists of 15 different offense categories. Each category is classified into suboffense category ranging between 3 and 58 subcrime categories. We have to consider the offense categories that have less than 20 subcategories.

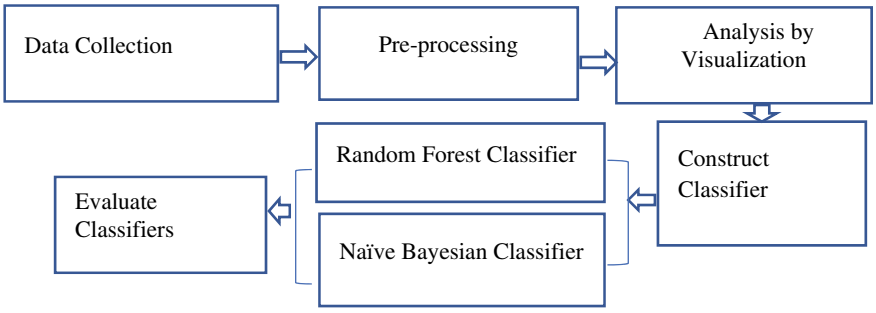


Fig. 6.1 Crime data visualization and prediction process

Analysis by Visualization
One important aspect of this study is to design methods of analysis and interpretation of crime data through visual analytics. To determine frequent prone areas of crime, offense category, period of occurrences, neighborhood locations that are victims of similar type of crimes, data visualizations are applied. Visualization provides a quick summary of huge data.

Construct Classifier
The work considers two popular machine learning approaches *probabilistic* and *ensemble methods. Probabilistic:* Given an observation of input, a probability distribution over a set of classes the **probabilistic classifier** can predict the output class. Bayes theorem $P(A|B) = (P(B|A)P(A))/P(B)$, the probability of A is predicted given the occurrence of B in dataset with strong independence among the features of dataset, the Naive Bayes approach is designed. This work uses Naive Bayes model to predict the offense category for Crime dataset. Naive Bayes is simple and fast classifier, highly scalable, requires less training data, can be used for binary, multi-class classification problems. Naive bayes is often "good enough" in a lot of real-world applications. When data is dynamic and keeps changing, Naïve Bayes can adapt quickly to the changes. *Ensemble*: These methods uses multiple learning algorithms to obtain better prediction than obtained by any single learning model. We employed random forest in this work. Random forest method is a decision forest method applied to Classification and Regression tasks. The method constructs multiple Decision trees at learning stage and outputs a model that has accurate many a times. With random forest approach, overfitting is reduced by which the prediction accuracy improves. Multiple trees reduce the chance of stumbling across a classifier that doesn't perform well. In case of unbalanced datasets, random forest has balancing error in class population. It has capabilities to compute similarities in the data and identify outliers. Thus, it can be extended to unlabeled data, leading to unsupervised learning, data views and outlier detection.

Section 6.4 deals with experimental approach that discusses the analysis of dataset using visualization and crime patterns using prediction models. Section 6.5 discusses evaluation of the prediction models and in Sect. 6.6 conclusions.

6.4 Experimental Approach

We explore R language for data preprocessing, visualization, classification. Shiny Package was used for design of Webpage, embed standalone apps or to build dashboards. The significance of Shiny is it provides extension with CSS, HTML widgets and JavaScript Actions. Shinythemes package provides a flexibility to change the overall appearance of a webpage. This work explored date attribute for knowing the frequent days of crimes or traffic activities, removal of null valued records, selecting significant attributes for classification. Lubridate package is used extensively for manipulating date attribute to better understand the crime occurrence patterns based

on day, month, time in terms of hours of the day. In Data Visualization phase, we extensively used Plotly and ggplot2 and ggthemes packages for crime data analysis. Ggplot2 is a system for declaratively creating graphics, based on the grammar of graphics. Just provide the data, tell ggplot2 how to map variables to aesthetics, what graphical primitives to use, and it takes care of the details. It embodies a deep philosophy of visualization. For interactive data visualization with high-quality graphs online, Plotly R graphing library makes interactive, publication-quality graphs online.

6.4.1 Result Analysis by Visualization

Crime Data Analysis and Visualization is a systematic approach for identifying and analyzing patterns and trends in crime. The work here experimented by analyzing crimes through visualizing the prominent offense categories based on month in an year recorded more offenses of a particular category of offense, on which day of the week such offenses are more to happen and so on. Using the ggplot2 of R, we can visualize through map the offense category and the regions where it has occurred. Our observations from the data are similar type of offense categories take place in the neighborhood areas. We explored neighborhood regions and their crime nature.

Table 6.1 shows different offense category types and the no. of cases recorded during 2012–17. It can be seen that 13 different categories of crime types with different sub-category types in each category. For our work, we considered first five crime types for analysis using visualization.

Table 6.1 Offense category-wise crimes recorded

Sl. No.	Offense_category_id	No. of cases
1	Traffic-accident	104,739
2	Larceny	43,334
3	Public-disorder	41,269
4	Drug-alcohol	33,384
5	Burglary	28,354
6	Theft-from-motor-vehicle	23,085
7	Auto-theft	21,679
8	Other-crimes-against-persons	17,701
9	Aggravated-assault	8680
10	White-collar-crime	5944
11	Robbery	4994
12	Arson	498
13	Murder	169

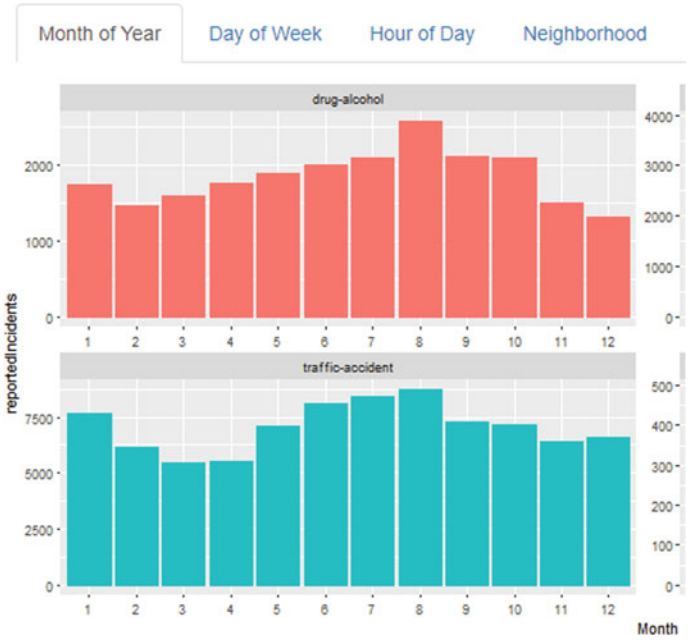


Fig. 6.2 Histogram of traffic accidents, drug crimes month wise

From Fig. 6.2, It is observed that in the month of August, drug-alcohol crimes were high and consistent throughout the year the crimes are above 1000 excluding the month of December which is slightly less. Figure 6.3 shows that these cases occurred in the mid of the week and are very less on weekends. It can be interpreted as people indulge in drug-alcohol crimes in the weekdays under the influence of friends or colleagues. Similarly, Fig. 6.4 explores public disorder, alcohol, burglary and larceny cases. Public disorder which constitutes gambling, criminal mischief, threaten to injure, harassment, disturbing peace is some subtypes of public disorder occurred consistently in all weekdays, but were at peek during June to August months, traffic accidents were at peek between June–August and quiet low in March and April. Clear weather may be one of the reasons for less accident cases in March–April. Larceny cases are consistent in weekdays, and more on weekends. Burglary crimes are more in April and consistent on all weekdays. Knowing the month and weekdays of peek crime types will help the authorities to take appropriate measures to address them.

Figure 6.5 depicts the hour of the day the crime taken place. It is observed that drug-alcohol cases are more in the evening hours and slowly reduced in the nights. Similarly, traffic accident in 3rd quadrant of the figure is during the office hours, slowly reduced toward night. Public disorder cases are more or less consistent throughout the day from 9 am onward. White-collar cases are during the office hours, and they are note minimum in the morning hours of the day.

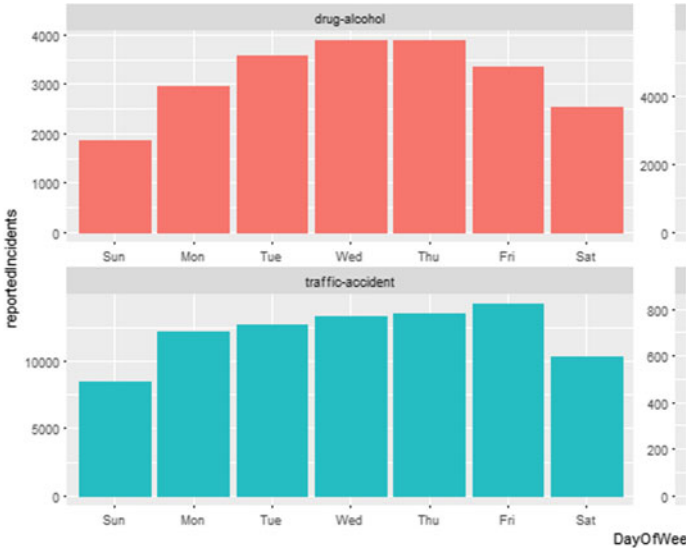


Fig. 6.3 Histogram of traffic accidents, drugs day of the week

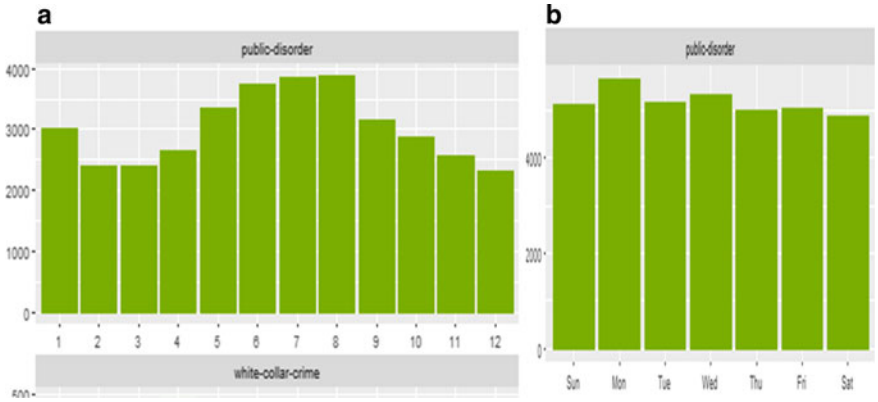


Fig. 6.4 **a** Public disorder crimes month of the year, **b** day of the week

Another type of analysis is identifying the neighborhood cities that are victims of similar type of crimes. In total, 80 neighborhood cities exist in dataset. Table 6.2 shows maximum and minimum count of crime for five neighborhood ids are listed.

It should also be noted for the five offense categories that this paper study shows that five-points is a place where public disorder, burglary and drug alcohol cases are very high and Wellshire, DIA and Indian-creek are the places of least occurred cases. Traffic accidents and Larceny cases were noticed many in Stapleton and less in Wellshire. This may end up in interpreting that Wellshire, Indian-Creek are the

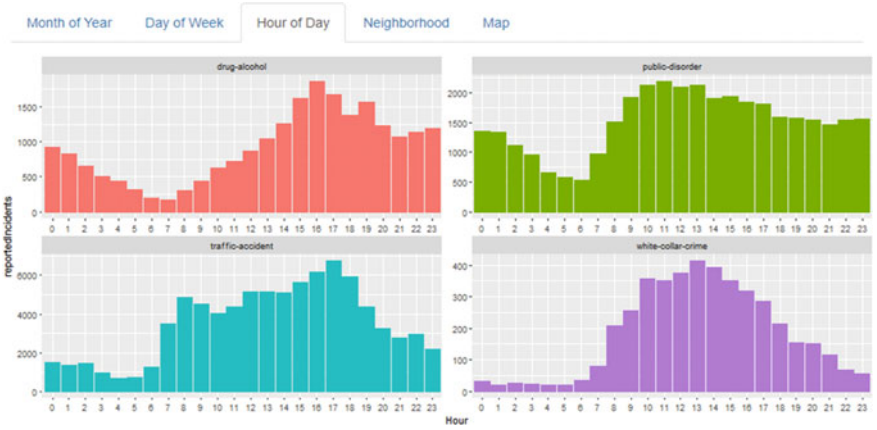


Fig. 6.5 Histogram of crime occurrence in the hour of day

Table 6.2 Neighborhoods maximum and minimum crime occurrences

Neighborhood Id	Maximum no. of crimes	Neighborhood Id	Minimum no. of crimes
Five-points	18,761	Rosedale	1300
Stapleton	14,463	Skyland	1178
Capitol-hill	12,636	Ry-club	928
Cbd	11,842	Wellshire	495
Baker	9981	Indian-creek	444

peaceful and safe places when compared to five-points and Stapleton.Visualization provides a broad range of analysis without knowing much about the dataset. Quick analysis and collective measures will reduce the crime.

6.4.2 Construction of Prediction Models

We have applied Naive Bayes and Random Forest algorithms on the dataset to predict the crime category. The test set contained 25 records. Train set contains 9 categories. The category 'others crimes against persons' has 58 subcategories; therefore, this category is excluded from the dataset. The dataset for prediction has 240,562 records out of which 168,393 were training data which is 70% of the dataset and 72,169 is the testing data.

Naïve Bayes model details of test and train data, construction and prediction are seen in Figs. 6.6, 6.7, 6.8 and 6.9.

Fig. 6.6 Levels in offense category in train set

```
Console E:/zareen/Crime_Data_DS_Alg/
> levels(train$OFFENSE_CATEGORY_ID)
[1] "aggravated-assault"
[2] "arson"
[3] "auto-theft"
[4] "burglary"
[5] "larceny"
[6] "murder"
[7] "robbery"
[8] "theft-from-motor-vehicle"
[9] "traffic-accident"
>
```

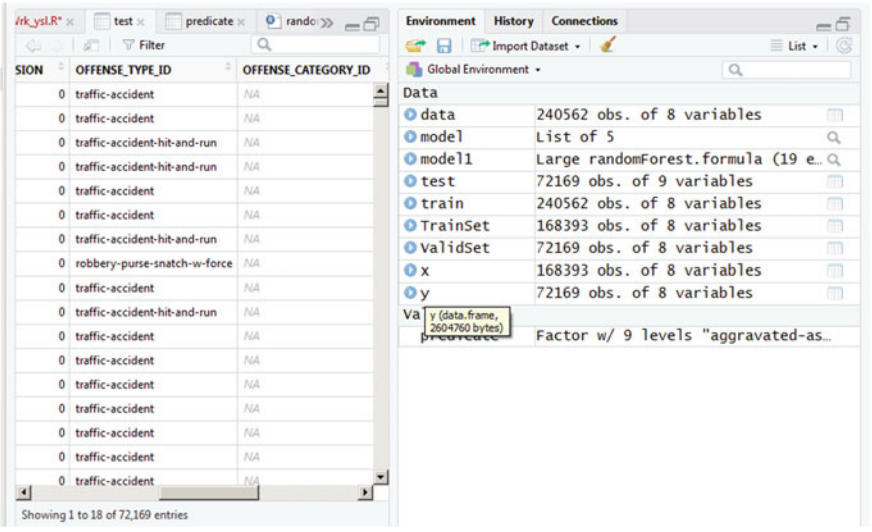


Fig. 6.7 View of environment and test set

Figure 6.6 depicts the offense categories in the dataset. Figure 6.7 depicts the environment variables and test set details. Figure 6.8 shows the Naïve Bayes model construction and Fig. 6.9 prediction on test set.

Random Forest model details of test and train data, construction and prediction are seen in Figs. 6.9, 6.10 and 6.11. Random forest algorithm works when the no. of sub categories are less than 53. Due to this reason, we removed other crimes from people category records from the dataset. Figure 6.9 depicts the Random forest model, train set details. Figure 6.10 shows the confusion matrix of random forest offense categories predictions and Fig. 6.11 prediction on test set and compares with that of Naïve Bayes model.

```
> model<-naiveBayes(train$OFFENSE_CATEGORY_ID~.,data = train)
> model

Naive Bayes Classifier for Discrete Predictors

Call:
naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:
Y
      aggravated-assault                   arson              auto-theft
            0.0360821742            0.0020701524            0.0959627871
                burglary                 larceny                  murder
            0.0901181400            0.1801365137            0.0007025216
                 robbery theft-from-motor-vehicle         traffic-accident
            0.0207597210            0.1387750351            0.4353929548

Conditional probabilities:
                          offense
Y                             [,1]      [,2]
  aggravated-assault       239534.4 108734.11
  arson                    250242.7  98530.74
```

Fig. 6.8 Naïve Bayes model training

```
> model1 <- randomForest(x$OFFENSE_CATEGORY_ID ~ ., data = x, importance = TRUE)
> model1

Call:
 randomForest(formula = x$OFFENSE_CATEGORY_ID ~ ., data = x, importance = TRUE)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 2

        OOB estimate of  error rate: 0%
```

Fig. 6.9 Random forest model construction

```
Confusion matrix:
                          aggravated-assault arson auto-theft burglary larceny murder robbery
aggravated-assault                     6040     0          0        0       0      0       0
arson                                     0   343          0        0       0      0       0
auto-theft                                0     0      16196        0       0      0       0
burglary                                  0     0          0    15179       0      0       0
larceny                                   0     0          0        0   30386      0       0
murder                                    0     0          0        0       0    108       0
robbery                                   0     0          0        0       0      0    3485
theft-from-motor-vehicle                  0     0          0        0       0      0       0
traffic-accident                          0     0          0        0       0      0       0
                          theft-from-motor-vehicle traffic-accident class.error
aggravated-assault                               0                0           0
arson                                            0                0           0
auto-theft                                       0                0           0
burglary                                         0                0           0
larceny                                          0                0           0
murder                                           0                0           0
robbery                                          0                0           0
theft-from-motor-vehicle                     23371                0           0
traffic-accident                                 0            73285           0
> predicate <- predict(model1, ValidSet)
```

Fig. 6.10 Confusion matrix of random forest offense category prediction

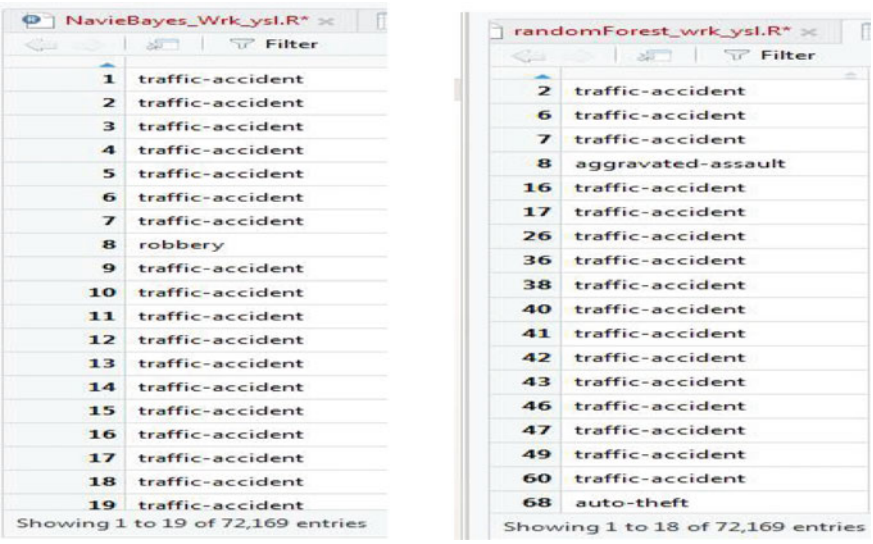


Fig. 6.11 Predicted values using naïve bayes and random forest

Table 6.3 Naive Bayes wrongly classified classes

Correct class	Wrongly classified
Auto theft	Theft-from-motor-vehicle
Burglary	Theft-from-motor-vehicle, aggravated assault, auto-theft
Larceny	Aggravated assault, robbery
Murder	Theft-from-motor-vehicle

6.5 Evaluation

The observations from prediction models are random forest generated 500 trees, and its accuracy is 100%. It could correctly predict the classes of Offense Category for the dataset of about 72,169 records. The problem with the model is it took more time to build the model. Prediction was fast. Whereas Naïve Bayes algorithm took less time to build the model, but its accuracy is given 98.42%. It wrongly classified 1134 records. Its Prediction time is slow (Table 6.3).

6.6 Conclusions

The objective of this work is to study the impact of data visualization in analysis and interpretation of huge data. In this process, we demonstrated how data and

visual analytics together will assess the characteristics of Crime Data. Using visual analytics, we could summarize that traffic accidents are more common offenses, apart from burglary and public disorders. We have also noticed that there are peaceful areas and crime-prone areas by identifying commonly occurring neighborhood areas. Visualization presents a quick summary of crime types and its network. Neighborhood analysis of offense category summarizes the crime-prone areas. Crime data analysis can help law enforcement officers to take collective measures increase level of crime prevention. Our prediction models are working quite well. But random forest has a limitation that more than 53 subcategories are not accepted. In our future work, we want to look in this direction. We wish to work with more advanced versions of machine learning algorithms, with huge data. Using Spark, we want to study this data more precisely.

References

1. Bogomolov, B., Lepri, J., Staiano, N., Oliver, F., Pianesi, Pentland, A.: Once upon a crime: towards crime prediction from demographics and mobile data. In: Proceedings of the 16th International Conference on Multimodal Interaction, pp. 427–434 (2014)
2. Iqbal, R., Murad, M.A.A., Mustapha, A., Shariat Panahy, P.H., Khanahmadliravi, N.: An experimental study of classification algorithms for crime prediction. Indian J. Sci. Technol. **6**(3) 4219–4225 (2013)
3. Chen, H., Chung, W., Xu, J.J., Wang, G., Qin, Y., Chau, M.: Crime data mining: a general framework and some examples. IEEE Comput. **37**(4), 50–56 (2004)
4. Beshah, T., Hill, S.: Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia. In: Proceeding of Artificial Intelligence for Developments (AID 2010), pp. 14–19 (2010)
5. Zhang, Q., Yuan, P., Zhou, Q., Yang, Z.: Mixed spatial-temporal characteristics based crime hot spots prediction. In: IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Nanchang, China (2016)
6. Sri Lalitha, Y., Govardhan, A.: Semantic framework for text clustering with neighbors. In: ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of CSI, Volume II, Advances in Intelligent Systems and Computing 249, pp. 261–271. © Springer International Publishing, Switzerland (2013). ISBN: 978-3-319-03095-1
7. Mahmud, N., Ibn Zinnah, K., Ar Rahman, Y., Ahmed, N.: CRIMECAST: a crime prediction and strategy direction service. IEEE 19th International Conference on Computer and Information Technology. Dhaka, Bangladesh (2016)
8. Lin, Y.L., Yu, L.C., Chen, T.Y.: Using machine learning to assist crime prevention. In: IEEE 6th International Congress on Advanced Applied Informatics (IIAI-AAI). Hamamatsu, Japan (2017)
9. Grover, V., Adderley, R., Bramer, M.: Review of current crime prediction techniques. In: International Conference on Innovative Techniques and Applications of Artificial Intelligence, pp. 233–237. Springer, London (2007)
10. McClendon, L., Meghanathan, N.: Using machine learning algorithms to analyze crime data. Mach. Learn. Appl. Intl. J. (MLAIJ) **2**(1) (2015)
11. Budur, E., Lee, S., Kong, V.S.: Structural analysis of criminal network and predicting hidden links using machine learning. arXiv:1507.05739 (2015)
12. Chainey, S., Ratcliffe, J.: GIS and Crime Mapping. Wiley, USA (2015)
13. Kim, S., Joshi, P., Kalsi, P.S., Taheri, P.: Crime analysis using machine learning. In: IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (2018)

Chapter 7
Automatic Reminder List Generation to Assist People with Alzheimer's



Shireesha Gundeti and Bhageshwari Ratkal

Abstract No matter how well people set up to-do lists and calendars, they usually can't get things done unless they have a reliable way of reminding themselves to do them. However hard they try to remember them all, it is very common for people to forget an appointment with a doctor, or to wish a friend on birthday. And moreover, as people grow older, their brains change and they may have problems related to memory, every once in a while. These details may include simple things like not remembering when to take medicines or important details like forgetting their own house address. An application that can automatically generate a reminder list from people's conversation and remind them the appropriate task on appropriate time would definitely help people in avoiding such problems. The proposed idea would avoid the need of people themselves typing in the task and setting the time for the reminder, manually, as this application would take the person's casual conversation as input and would generate the reminder list automatically based on the related information extracted from the conversation. And based on the automatically generated reminder list which would be connected to one's calendar, the user would receive a notification when the clock ticks the time on a particular date as set in the reminder, thus helping the user to remember and do all the tasks at the required time.

7.1 Introduction

People always forget to get their things done, no matter how well they set up their to-do lists or calendars. It's very natural for even a young person to forget to catch up on a business meeting or to attend a family get-together. Usually, people tend to forget those events or tasks that are not their habits. If we suppose that all the events or tasks in a typical human life be classified into two, then they would be

S. Gundeti (✉) · B. Ratkal
Department of Computer Science and Engineering, G. Narayanamma Institute of Technology and Science, Shaikpet, Hyderabad, India

B. Ratkal
e-mail: bhagya@ratkal.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_7

habitual events/tasks and occasional or random events/tasks [2]. Habitual events include events such as brushing teeth, bathing, having food, and a regular person would never forget or misdo such tasks. Whereas occasional or random events include meetings, appointments, birthdays, etc. And forgetting a couple of events of this kind is a tendency that is often seen even in young and healthy human beings. Also, this tendency is observed to grow, as people grow older. And, elderly people are more vulnerable to being attacked by Alzheimer's. Alzheimer's disease causes problems with memory, thinking, and behavior [1]. As the disease progresses, a person with Alzheimer's disease will develop severe memory impairment and lose the ability to carry out everyday tasks.

So, forgetting things is an absolute problem that almost every human being on earth is facing. And a reliable way of reminding people, about their events or tasks is all they need. To address this issue, many of the reminder setting apps have been introduced into the market. But the major drawback in almost all of the existing apps is that they require the users to set the reminders manually. This itself is a tedious and boring task; it also encompasses another risk of the user forgetting to set the reminder.

Now, to address this drawback, an application named Automatic Generation of Reminder List (AGRL), based on Semantic Analysis is developed that can automatically set reminders based on the conversations that people make, with each other, regarding any event or task [1]. That is, if the user is discussing, with someone, about some event that is going to be held in future, then this conversation would be recorded by the application and based on the conversation, the user would be reminded of the event on a particular date at a particular time as mentioned by the user in the conversation. This solution avoids the necessity of manual setting of reminders and thus helps people to not miss out on events. Especially people affected with Alzheimer's can get greater help from this application and can lead an easier routine life.

7.2 Existing Systems

Most of the available reminder apps have good reviews from their users and are used by a good portion of the population. Almost all of these apps have an easy-to-use interface, voice-enabled interfaces for setting reminders, and many more options for the users so that they won't miss out on their events [3].

To discuss in detail, if a user using any of these apps wants to set a reminder, the user should type the title or something about the event that can remind the user of the event, the date of the event or the task, and time to be specific. Or there is also another facility in a few of the apps available now, which provides a voice recording option so that the user can record the title or name of the event instead of typing it in. This is obviously a great help to people who are busy and don't have time to type in the title or to people who are not that comfortable with typing on the phone. But the point to be noted here is that even though entering the title is made easier, the task of

setting the reminder which includes selecting the date, time if it should repeat when to repeat and the number of times it should repeat, is still left to the user to do.

For the whole task of setting the reminder to be automated, it requires the automation of the following subtasks:

- Entering the title or name of the event
- Choosing the date
- Setting the time

So, the main objective of the proposed application is to automate the process of setting the reminders which include the above-mentioned subtasks.

7.3 Proposed System—Automatic Generation of Reminder List (AGRL)

The working of the proposed system is depicted in Fig. 7.1.

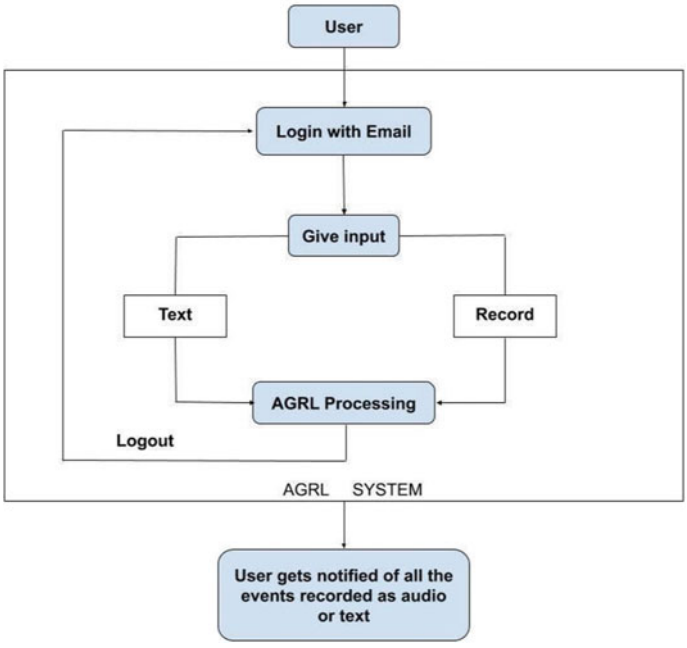


Fig. 7.1 AGRL Architecture

7.3.1 User Input

Once the user has logged into the application using a valid email-id, the user can give the input in either of the following two ways.

Text: The user can simply type in the event, time, and date, maybe, in the form of a simple sentence.
Record: Or the user can choose to record the conversation which may include information about some events and the dates when they might occur in the future.

7.3.2 AGRL Processing

The input might be in the form of text or audio, either way, it requires a lot of processing to extract the required task or event name and the associated time when it might take place. The methodology involved in this processing is elaborated in the section IV.

7.3.3 The User Gets Notified of All the Events Recorded as Audio or Text

Once the processing of the given text or audio is completed and the extracted reminders are connected with the user's Google calendar, the user will be reminded of all those events that were typed in by the user or those that were mentioned by the user in the conversation that was recorded in the application. At the time of the reminder, the user may receive a notification, according to the user's notification settings in the Google calendar app.

7.4 AGRL Processing

The steps followed in AGRL processing are depicted in Fig. 7.2. Let the AGRL processing part be classified into the following three sections.

7.4.1 Text—Preprocessing

This step comes into action, only when the user gives the input in the form of text. The text preprocessing includes only one step namely sentence tokenization [6]. And this step assumes that the user would stick to the following grammar conventions:

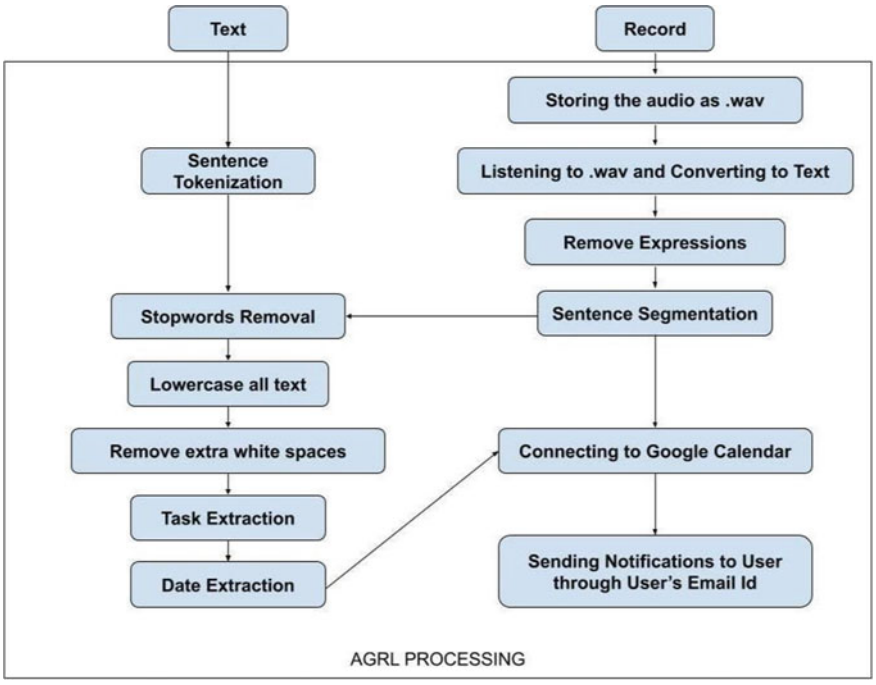


Fig. 7.2 AGRL processing

- Use clear and correctly formed sentences in English while typing in the text
- Use correct punctuation such as full stops(.), question marks(?), to indicate the end of the sentences, in case of more than one sentences
- Do not use chat language or short-cuts

The process involved in this step is as follows.

Sentence Tokenization

The text given by the user can be a single sentence or a paragraph. If the input is a single sentence, it goes unchanged in this step of preprocessing. But if the input contains more than one sentence or a paragraph, then it needs to be sentence tokenized, i.e., the given paragraph needs to be split into sentences at the specified sentence separators, which in most cases is a full stop(.).

7.4.2 Record—Preprocessing

If the user gives the input in the form of audio, only then the processing steps to be mentioned under this section would come into play. The steps involved are as follows.

Storing the Audio as .wav
The speech or the conversation recorded by the user would be initially stored as an audio file with the.wav extension. The audio content in this file will be processed in further steps.

Listening to .wav and Converting to Text
Here a speech recognition is used to listen to the audio content in the .wav file and convert it to text. Only if the speech or conversation recorded is clear, audible, and loud enough, the converted text would be correct as per user and the following results would be as expected.

Remove Expressions
While recording the speech or conversation, if the user or the other speakers in the conversation show some kind of expressions such as laugh or smile or cry or grin, they will also get recorded in the audio and would get converted to text as well. But these expressions would add no value to the text related to the reminders and hence these should be removed or neglected before proceeding to extract reminders related text. To do so, when some of the speech-to-text conversions containing expressions are noticed, it is observed that when the sound associated with the expressions is converted to text, they are enclosed within some of the special expressions such as [] or {} or **, etc. For example, '[laughing]', '{crying}', *grins*, etc. Using the Re module in Python, the expressions are identified in the text and eliminated.

Sentence Segmentation
Unlike the case when the input is given as text, when the input is given as audio and later converted to text, there would be no sentence separators specifying the end of each sentence and hence the process used in sentence tokenization would not be appropriate to use here. In order to address this situation, sentence segmenter is used.

7.4.3 Text and Record—Processing

After the input is processed as per the process mentioned in the above sections, the resulting text is further processed through the following steps. Note that the steps to be mentioned are commonly performed in both cases of input, text, or audio.

Stopwords Removal
All the not so important words in each sentence need to be removed as they sometimes might add inaccuracy to the final output obtained after processing the input.

Lowercase all Text
Lower casing all the text is done to avoid having multiple copies of the same words. For example, while analyzing the words or pattern matching, 'Analytics' and 'analytics' will be taken as different words, if not lowercased before processing.

Remove Extra White Spaces
Extra white spaces, if present in the given text, also should be removed, so that all the words in the text are separated by single spaces. This step is needed because sometimes extra spaces might also cause problems while pattern matching.

Task Extraction
The task extraction is carried out using Parts-of-Speech Tagging and Chunking [5]. The sub-steps involved in this step are as follows:

- Tokenize text: Each sentence is tokenized into words.
- POS_Tag is applied to each of the words obtained in the above step.
- POS Tagging is combined with Regular Expressions to form and extract chunks of noun and verb phrases.

Basically, a chunk grammar needs to be defined to specify the rules that indicate how the sentences should be chunked [4]. And the grammar that was defined for the task at hand is as follows.

grammar = "NP:{<VB.*>?<VBP>?<RB>?<PRP.*>?<IN>?<DT>?<VBZ>? <VBN>?<POS> ?<NN>?<VBZ>?<JJ.*>*<NN.*>+<VBZ>?<TO>?<VB>?}"

The above grammar specifies that the NP chunk should be formed whenever the chunker finds an optional verb type, followed by an optional verb in present tense without third person singular, followed by an optional adverb, followed by an optional personal pronoun, followed by an optional preposition, followed by an optional determiner, followed by an optional verb in the present tense with third-person singular, followed by an optional verb past participle, followed by an optional possessive ending, followed by an optional noun, followed by an optional verb in the present tense with third-person singular, followed by any number of adjectives, followed by one or more number of nouns, followed by an optional verb in the present tense with third-person singular, followed by an optional infinite marker, followed by an optional verb.

The chunks that are formed in each sentence are ranked based on the summation of their individual POS ranks set as follows.

POS_rank = {"DT":2, "IN":2, "RB":3, "PR":3, "VB":6, "JJ":5, "NN":5, "PO":5,"TO":1}

Among the chunks formed, one with the maximum chunk rank would be considered as the required task name that can remind the user of the event.

Date Extraction
To extract the date and time relevant to the reminder that needs to be reminded, different cases in which the date and time can be specified in a sentence are all analyzed. And for all the cases that were analyzed, a regular expression pattern was

constructed, so that when a date-time string matching the corresponding pattern is found, it would trigger the required processing needed in order to extract the date and time from the given sentence.

The regular expression in the python programming language is a method used for matching text patterns. The 're' module in Python provides regular expression support.

7.5 Discussions on Results

The main aim of the application, which was to assist the people affected with Alzheimer's to remember things that can be reminded, is achieved. The performance of the application is satisfactory in terms of accuracy, security, speed, usability, etc.

To test the accuracy of the model used in the application, more than 1000 sentences, that may contain information related to reminders, were collected from some random online websites [7]. The labels and the datetime information, if present, were manually extracted from each of those sentences. And all this information is organized in the form of a dataset containing 3 attributes; sent, label, datetime. This dataset is shown in Fig. 7.3.

The values that these attributes can take are as follows.

sent: The sentence that may contain information related to reminders. In other words, it can be any sentence that may or may not indicate a reminder.
label: The task or event name related to the reminder, if present in the sentence. The values that this attribute would take can be categorized into two.

- if task name or event name appears in the sentence → task or event name
- else → Not Found

	sent	label	datetime
0	Set a reminder on date 23rd November 2016	Not Found	23-11-2016 00:00
1	remind me 6 pm today evening	Not Found	06-04-2020 18:00
2	Remind me at 28 December for recharge	recharge	28-12-2020 00:00
3	remind me at 7pm on 8 Jan	on 8 Jan	08-01-2020 19:00
4	Set a reminder on 4th Dec of going to meet son...	meet sonal miss	04-12-2020 00:00
...	...	...	...
1177	Remind me to go to Eat Panipuri at 5:00 PM today	go to Eat Panipuri	04-06-2020 17:00
1178	Remind me to order garnier roll on 1 January at 1	order garnier roll	01-01-2020 13:00
1179	Please remind the same on 13th November	Not Found	NaN
1180	Hi I want alarm with tone saying time is 3.30a...	3.30am...it's time booja	04-06-2020 15:30
1181	Hey remind me to mail the scanned copy of canc...	mail the scanned copy of cancelled check to Vi...	04-06-2020 16:00

1182 rows × 3 columns

Fig. 7.3 Testing dataset

	sent	label	datetime	label_testing	datetime_testing
0	Set a reminder on date 23rd November 2016	Not Found	23-11-2016 00:00	Set a	23-11-2016 00:00
1	remind me 6 pm today evening	Not Found	06-04-2020 18:00	Not Found	06-04-2020 18:00
2	Remind me at 28 December for recharge	recharge	28-12-2020 00:00	for recharge	28-12-2020 00:00
3	remind me at 7pm on 8 Jan	on 8 Jan	08-01-2020 19:00	Not Found	08-01-2020 19:00
4	Set a reminder on 4th Dec of going to meet son...	meet sonal miss	04-12-2020 00:00	meet sonal miss	04-12-2020 00:00
...	...	...	...	...	...
1177	Remind me to go to Eat Panipuri at 5:00 PM today	go to Eat Panipuri	04-06-2020 17:00	Eat Panipuri	06-04-2020 17:00
1178	Remind me to order garnier roll on 1 January at 1	order garnier roll	01-01-2020 13:00	order garnier roll	01-01-2001 00:00
1179	Please remind the same on 13th November	Not Found	NaN	Not Found	13-11-2020 00:00
1180	Hi I want alarm with tone saying time is 3.30a...	3.30am...it's time booja	04-06-2020 15:30	saying time is	06-04-2020 03:00
1181	Hey remind me to mail the scanned copy of canc...	mail the scanned copy of cancelled check to Vi...	04-06-2020 16:00	of cancelled check to	06-04-2020 16:00

1182 rows × 5 columns

Fig. 7.4 After adding the columns generated by the AGRL model

datetime: The date and time information related to the reminder, if present in the sentence. The values that this attribute would take can be categorized into two.

- if the date and time information appear in the sentence → datetime
- else → Nan (empty)

Once this dataset was ready with around 1100 records, the model used in the application was executed giving the sentences, from this dataset, as input to give the label_testing and datetime_testing values as output. After adding this output as another two columns in the previous dataset, it looked as in Fig. 7.4.

Here, the extra attributes indicate the following:

label_testing: The task name or event name extracted from the sentence, by the model constructed.

datetime_testing: The datetime information extracted from the sentence by the model used in the application.

Now, the accuracies of individual results of extraction of labels and datetime were estimated, and then the accuracy of the combined output was calculated using the following formulas.

$$\textbf{Label\_Accuracy} = \frac{\text{Number of Correctly Estimated Labels}}{\text{Total number of records}} \times 100$$

$$\text{where, Correctly Estimated Labels} = \begin{cases} 1; \text{if label} = \text{label\_testing} \\ 0; \text{otherwise} \end{cases} \tag{7.1}$$

$$\textbf{Datetime\_Accuracy} = \frac{\text{Number of Correctly Estimated Datetime}}{\text{Total number of records}} \times 100$$

$$\text{where, Correctly Estimated Datetime} = \begin{cases} 1; \text{ if datetime} = \text{datetime\_testing} \\ 0; \text{ otherwise} \end{cases} \tag{7.2}$$

$$\textbf{Label\_Datetime\_Accuracy} = \frac{\text{Number of Correctly Estimated Labels and Datetime}}{\text{Total number of records}} \times 100$$

$$\text{where, Correctly Estimated Labels and Datetime} = \begin{cases} 1; \text{ if label} + \text{datetime} = \text{label\_testing} + \text{datetime\_testing} \\ 0; \text{ otherwise} \end{cases} \tag{7.3}$$

Using the Eqs. (7.1), (7.2), and (7.3), the accuracies were estimated which were found to be 71% for label extraction, 84% for Datetime extraction, and 60% for the combined extraction. The accuracies obtained are pictorially depicted in Figs. 7.5, 7.6, and 7.7.

In Fig. 7.5, it is clearly shown that 71% of the labels were correctly extracted. And this accuracy was calculated based on the following assumptions.

- It is assumed that extracting something about the actual reminder is always better than not extracting anything. So, even if the label value in the dataset is 'Not

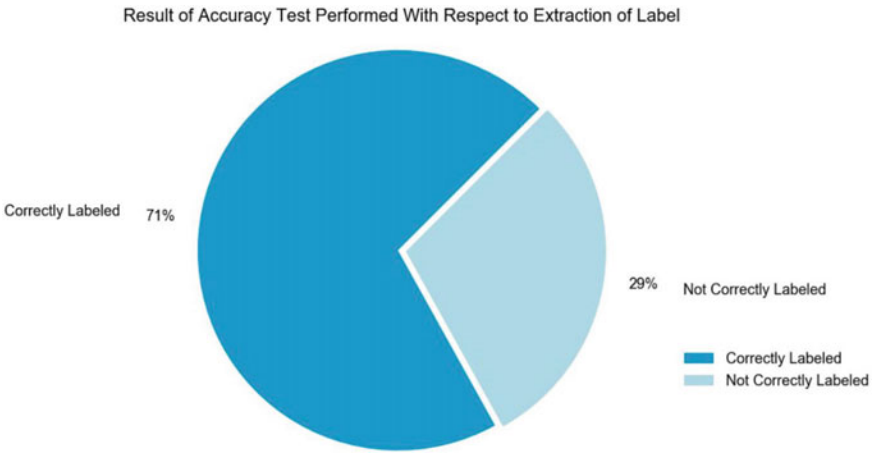


Fig. 7.5 Accuracy with regard to extraction of labels

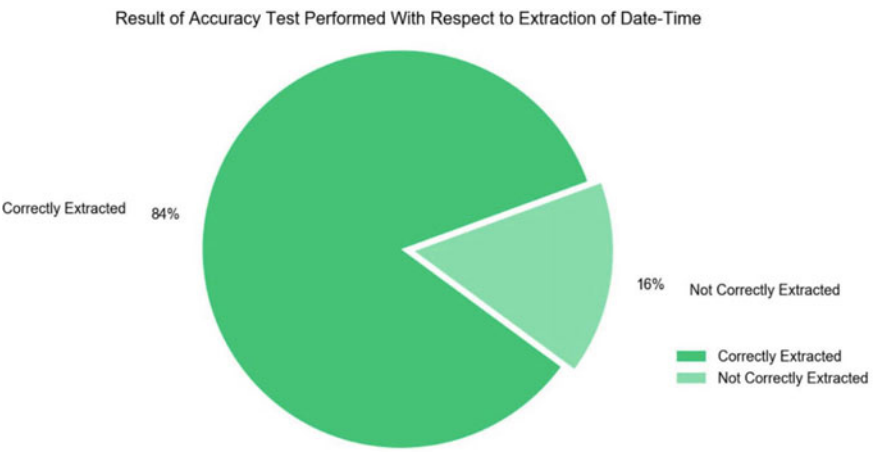


Fig. 7.6 Accuracy with regard to extraction of date–time

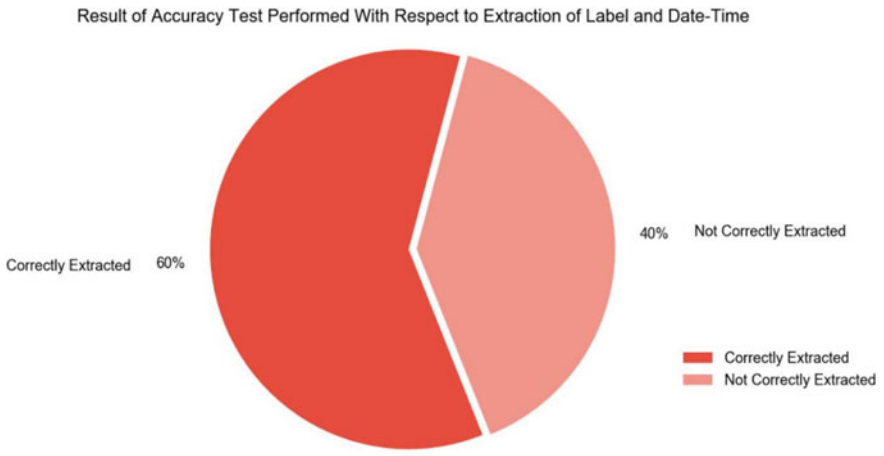


Fig. 7.7 Accuracy with regard to extraction of label and date–time

Found' and the label extracted using the application found to be some word from the input sentence, it is still considered as 'Correctly Labeled'.

- It is assumed that even a single word related to the event, to be reminded, is always enough to remind the user of the event. So, even if a single word in the label extracted using the application matches with that in the dataset, it is considered as 'Correctly Labeled'.

Figure 7.6 depicts the pie diagram, showing the percentages of both correctly and wrongly estimated dates and times related to the reminders. Unlike in the extraction of labels, no assumptions or exemptions are made in the extraction of dates and times, because a word predicted wrong can still remind the user of the event, but if the user is reminded of the event late, by even 5 min, is not admissible.

Here, while calculating the accuracy, the datetime from the dataset is exactly compared to that extracted using the application.

Figure 7.7**,** shows that about 60% of the extracted labels and datetime values are correct while 40% of them are wrongly estimated. Here, instead of separately extracting and comparing the results with those values in the dataset, this is to test the accuracy of the combined result extracted from the given input sentences. The steps followed in calculating this accuracy are based on the below algorithm.

- check if datetime from dataset matches exactly to the extracted datetime_testing
- If yes:
- check if label = = 'Not Found' or label_testing ! = 'Not Found'
- If yes: consider as Correctly Extracted
- If no: consider as Not Correctly Extracted

The bar graph in Fig. 7.8**,** has the three different Accuracy Tests Conducted on Extraction of Labels and Datetime on the X-axis and the Accuracy, obtained in each

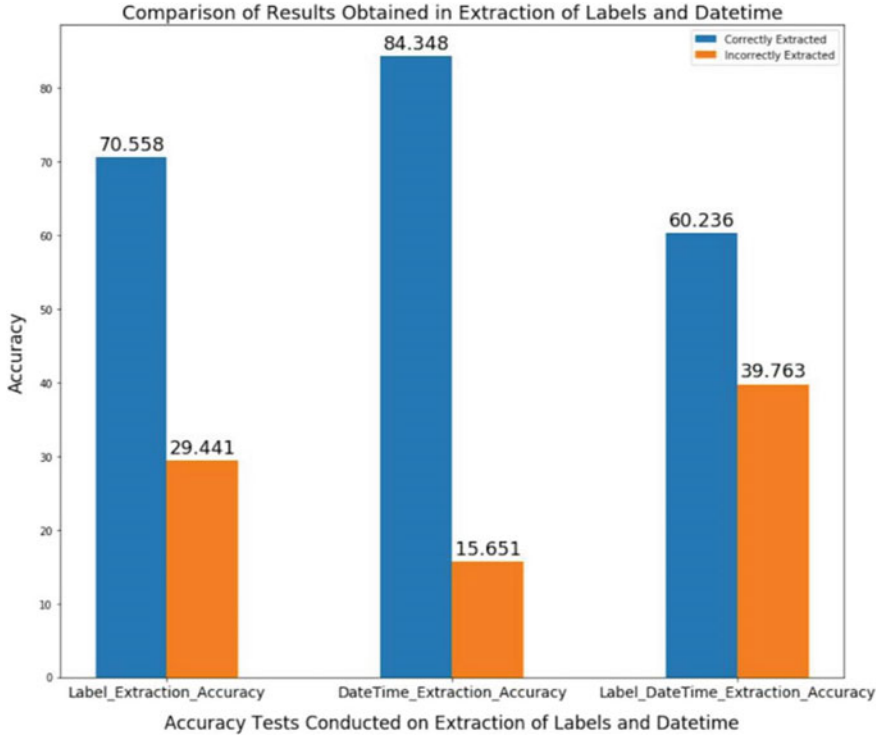


Fig. 7.8 Comparison of results obtained in the extraction of labels and datetime-I

test, on the Y-axis. It shows the comparison of percentages of the correctly extracted and incorrectly extracted results in each of the three test cases.

7.6 Conclusion

Using the proposed idea, any person suffering with Alzheimer's would now be able to attend almost all their events or tasks without fail. The users would now be able to avoid all the repetitious work involved in setting reminders.

References

1. https://www.mayoclinic.org/diseases-conditions/alzheimers-disease/symptoms-causes/syc-20350447
2. https://www.lifehack.org/articles/featured/back-to-basics-reminders.html

3. https://www.gadgetsnow.com/apps/6-best-reminder-apps-for-android-and-ios/articleshow/64684028.cms
4. https://www.guru99.com/pos-tagging-chunking-nltk.html#:~:text=Chunking%20is%20used%20to%20add,of%20more%20than%20one%20level
5. Varghese, N.,Punithavalli, M.: Lexical and semantic analysis of sacred texts using machine learning and natural language processing. Int J Sci. Technol. Res. 8(12) (2019)
6. https://www.nltk.org
7. https://www.researchgate.net/publication/330901752_Latent_Semantic_Analysis_Methods/link/5c5ab75245851582c3d1bdb2/download

Chapter 8
Recognizing the Faces from Surveillance Video



T. Shreekumar, K. Karunakara, H. Manjunath, B. Ravinarayana, and D. R. Annappa Swami

Abstract The face recognition (FR) is getting extremely well known in the field of biometric due to its unique nature. In spite of the fact that the progression of FR advancement has accomplished a moderate level of improvement with deep learning strategies, still there exists some disadvantage, for example, illumination variation, change in pose angle, occlusion, masked face, and so on. The deep learning strategies will be productive only if the learning database contains a large number of images for training purpose. It is a known fact that, a large number of images are very difficult to get always, especially from village people. In order to overcome all the discussed problems, this paper proposes two different FR approaches which functions admirably with the database even with few training images. The two methods proposed are (i) face recognition with local linear regression (LLR) using a trained artificial neural network (ANN), (ii) FR with support vector machine (SVM) by utilizing the well-known particle swarm optimization (PSO) technique for optimizing the SVM kernel parameters. The performance of both the techniques is measured in terms of statistical measures and the computation speed. The performance of the proposed methods is also compared with a very recent and similar active appearance model (AAM)-based FR. From the experiment and result analysis, it is clear that the proposed PSO-SVM-based FR yields a maximum accuracy of 95%, an average accuracy of 91.5% with a minimum time requirement of 0.03 s.

T. Shreekumar (✉) · H. Manjunath · B. Ravinarayana · D. R. A. Swami
Mangalore Institute of Technology and Engineering, Moodabidre, Mangalore, Karnataka, India

D. R. A. Swami
e-mail: annappaswamy@mite.ac.in

K. Karunakara
Sri Siddhartha Institute of Technology, SAHE, Tumakuru, Karnataka, India
e-mail: karunakarak@ssit.edu.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_8

8.1 Introduction

The requirement of accurate person identification system is very essential nowadays, in most of the government and commercial applications. The demand for the reliable user authorization methods is increasing because of the rapid developments in the fields of social networking, communication, and mobility. Biometric is a method of identifying the genuineness of the user, based on the physiological or behavioral traits. Biometric systems are now used in various applications like forensics, commercial, and other civilian applications. Face, signature, hand geometry, fingerprints, retinal scan, iris scan, hand vein, facial thermogram, voice, etc., are prominent ones in the biometric space. Biometric is getting popular starting from issuing an Adhaar card to gaining permission to enter a nation.

Face recognition forms the most trending technology nowadays. It forms the major way for identifying a person and also it becomes a better way for providing the much needed security for many kinds of applications. It forms the most reliable biometric due to its potential, when it is compared with other biometrics. This system has a long range of applications, both in the commercial as well as personal fields. The face recognition system uses many features from the face image which are used for recognizing the people. Although in the olden days, the face recognition equipments were using only geometrical features derived from the face directly, and they were not as accurate as they are now. Of late, due to the technological advancements, various methods which are used to identify the person are readily available. Nowadays, the face identification technology is used even for scam revealing and finding criminals too.

Face acknowledgment by utilizing an existing face image set verifies and confirms the individuals in the specified video stream or in the still image. [1–3]. It has an amorous application in security and authentication-related applications [4]. The entire face recognition methods can be divided into feature-based and appearance-based.

Feature-based method uses the geometric parameters from the face image for facial matching [7, 8] and appearance-based method uses intensity parameter [1, 17]. The face acknowledgment should be possible in the given picture also in the given video. The procedure of face recognition method is performed on the still image or even in a video streaming. Face detection and the face identification are the two stages present in a face recognition system [4]. In the face detection stage, the face is located in the video stream or in the still image. In the face detection stage, the face is located in the image or in the video stream. Then, the face identification algorithm utilizes the located face for identifying the individual from the image. Hence, both face detection and recognition algorithms are reasonably important in facial identification [13].

As it has been proved that motion enables identification of (well-known) faces when negation, inversion, or threshold operation is performed on images, video-based face recognition is preferred over using still image-based face recognition [3]. It has also been proved that animated faces compared to randomly reordered images from

the same collection are better recognized by humans. The video-based facial identification methods that consistently use both spatial and temporal information began only some years before and still require more investigation [9]. But the difficulty in video-based recognition is low frame quality [3, 11].

A good face recognition algorithm should overcome the bottlenecks like pose change, variations in illumination, aging, partial occlusion, sunglass, make-up, facial expression variation, change in hairstyle, scale variation, etc. Recognizing an image under severe illumination change and pose change is an extremely challenging problem in face recognition [12]. Illumination change can be handled by feature-based method easily and when the problem is with the pose, and the appearance-based method is predominantly used to recognize the faces. In appearance-based methods, the image of the individual is recognized from the image captured with a diverse in view angles [15, 16].

The face recognition has plenty of applications which include securing the mobile and computer applications, Internet communication, computer entertainment, verifying a person, securing a system using CCTV, and such others. Lately, the convolutional neural network-based face recognition has shown significant progress in the efficiency of face recognition task, but still the face recognition remains a challenging problem when the question is of recognizing a person from a poor quality video sequence or a noisy image or even from a group photos.

8.2 Literature Survey

The [17] introduces a face recognition method to identify the faces from variety of pose angles. This work is mainly focused on pose normalization and face recognition. Pose normalization is carried out with the help of local linear regression (LLR). The normalized face is then subjected to dimensionality reduction using principal component analysis (PCA). Then, the face image is recognized using Fisher's linear discriminant analysis (FLDA) from the reduced dimension. This approach has shown a good performance on Yale B Face Database, with 93.4% of accuracy.

The author [18] proposed a model for face identification using local binary pattern (LBP) and convolutional neural network (CNN). Initially, LBP features are extracted from the image and stored in a feature set. CNN training is carried out using this stored LBP features. The trained CNN is then used to identify the query image. From the experiment, author concluded that the input to CNN is an image then pixel level knowledge is extracted and the recognition rate is less. A substantial improvement in the recognition performance is observed when the LBP feature map is used as input to the CNN.

The work [19] has proposed a face acknowledgment framework to detect and identify the faces in still images and low resolution video sequence, by identifying the face images from the image gallery of known individuals, utilizing genetic algorithm and ant colony optimization techniques. In this method, initially, the pre-processing step is carried out on the input images, feature extraction is done using the ant

colony optimization, and finally, the recognition is carried out with genetic algorithm. This method is suitable in single and multi-threaded processing, real-time video processing. The proposed technique is robust to expression change and illumination change.

The face recognition method [20] has proposed a new method to identify the faces from a video sequence known as the 'analysis-by-synthesis framework' to overcome the variation in pose and illumination changes. This method works by tracking the entire video sequence to integrate the motion, lighting, and shape and to generate model from the gathered information. The information gathered from video tracking is used to estimate the illumination condition, pose alignment, and the structure of the face in each frame with the help of extended light source and multiple camera points. And then, corresponding to the illumination and pose conditions estimated from the video sequence, a 3D face model is generated. After that, the similarity measure is obtained by comparing the synthesized images and the probe video. This strategy can deal with circumstances where the face posture and lighting conditions in the training and testing information are totally disjointed.

In [21], an algorithm is proposed for recognizing the faces from a video sequence using the adaptive multiple face matchers. This method briefly populates the databases in order to capture the intraclass variations. The method uses active appearance model (AAM) to extract dynamic information about the facial poses in each frame and 3D face reconstruction model in order to factorize the same from the video sequence. The motion blur is estimated using discrete cosine transform (DCT). The experiment is carried on CMU's face-in-action (FIA) database and the result showed a good improvement when experimented with 204 subjects and with different face matchers.

In [22], a method is proposed for pose invariant face acknowledgment. Their face acknowledgment framework was insensitive to the changes in view direction and requires just one sample per person. The method used here is to find similarities of a face image against a set of face images of the same view, from the training set to establish a view invariant face acknowledgment.

The method [23] proposed a regression-based virtual frontal view generation in order to compensate pose variation. This frontal virtual view was then used to identify a person. They used local linear regression for virtual frontal view generation. They also tried with global linear regression (GLR) for virtual frontal view generation. The result indicated that the accuracy of LLR is maximum compared to the GLR-based virtual view generation. LLR uses a linear mapping function between the current view and the virtual frontal view. GLR uses the entire image and generates a virtual frontal image. The LLR divides the image into smaller blocks called, 'patches' with a fixed patch size of 20×20, or 10×10. They used dense sampling-based overlapped patch creation. Each patch is then subjected to LLR to get the linear mapping to the virtual frontal view. After the entire iteration of mapping for each patch, all the newly generated frontal patches are combined to form the virtual frontal face image. The experiment was carried out with CMU PIE face set and compared with eigen light-field method. The result indicates better compared to eigen light-field method.

The work [24] proposed a novel face recognition method for efficient face matching. Within this framework, the unprocessed greyscale face images of two individuals are compared using simple image processing filters. They successfully estimated the discrepancy in the illumination condition between the test image and the training set implicitly and used to weigh the contribution of query and the gallery representation. This filter-specific matching algorithm works offline by minimizing the online overhead. The experiment recorded the best performance, by correctly recognizing 97% of the individuals.

Face detection and identification from surveillance videos are quiet difficult because of image blur, illumination, pose variations, occlusion, and noise. A convolution neural network (CNN)-based recognition system is proposed in [25], which estimates the blur function from the difference between the artificially blurred images and the original image of the real-world video frames. Initially, the training data is formed using the original still images and the artificially blurred images of the same counterpart. The CNN is used to learn the blur function during the training phase, from the training data, made out of both the artificially blurred data and the still images. Secondly, in order to overcome the pose problem and occlusion, a Trunk-Branch Ensemble CNN (TBE-CNN) is proposed, which adequately separates the correlative data from the face images. Third, a triplet loss function is estimated to improve the effectiveness of feature representation of the extracted feature.

A facial recognition technique to extract the core region of face analysis, besides, all the while alter the facial posture and accomplish face alignment has been proposed in [26]. It can maintain a scheme distance from "mis-alignment disaster" amid the component matching process. The video face recognition system of little examples capacity had been built up to test the technique. The technique can create the face tests in the uncontrolled environment and upgrade the recognition rate of the face recognition system.

8.3 Face Recognition with LLR and ANN

In this approach, the face image is recognized with two different stages: (1) training section and (2) testing section. In the first section, an artificial neural network (ANN) is trained using the combined score [27] obtained from principal component analysis (PCA) and Fischer linear discriminant analysis (FLDA). In this method, the dimension reduction is done using PCA, feature vector is obtained using the FLDA, and then a combined score of PAC and FLDA is stored in a vector. The combined score of PCA and FLDA is then used to train the ANN. In the second section, the video of the person is obtained and normalized for illumination and pose variation problems. For illumination normalization, popular discrete cosine transform (DCT) and the pose problem are addressed with local linear regression method. The normalized frame is then subjected to get the combined score of PCA and LDA. The combined score is then used to recognize the face using already trained ANN. Figure 8.1 shows the structure of face recognition using local linear regression.

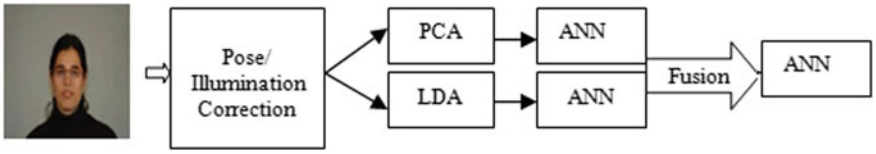


Fig. 8.1 PCA and LDA fusion using NN

8.4 Face Recognition with PSO and SVM

This method uses a popular and effective SVM to recognize the person from video frame. In this method, the SVM is trained with training database, and during the training phase, the SVM parameters are optimized with an optimization technique, particle swam optimization (PSO). Initially, the pre-processing of the image is carried out with adaptive median filter. In the next stage, the feature vector is obtained with active appearance model. Then, the SVM is used to recognize the face image. Figure 8.2 shows the structure of recognition system (Fig. 8.3).

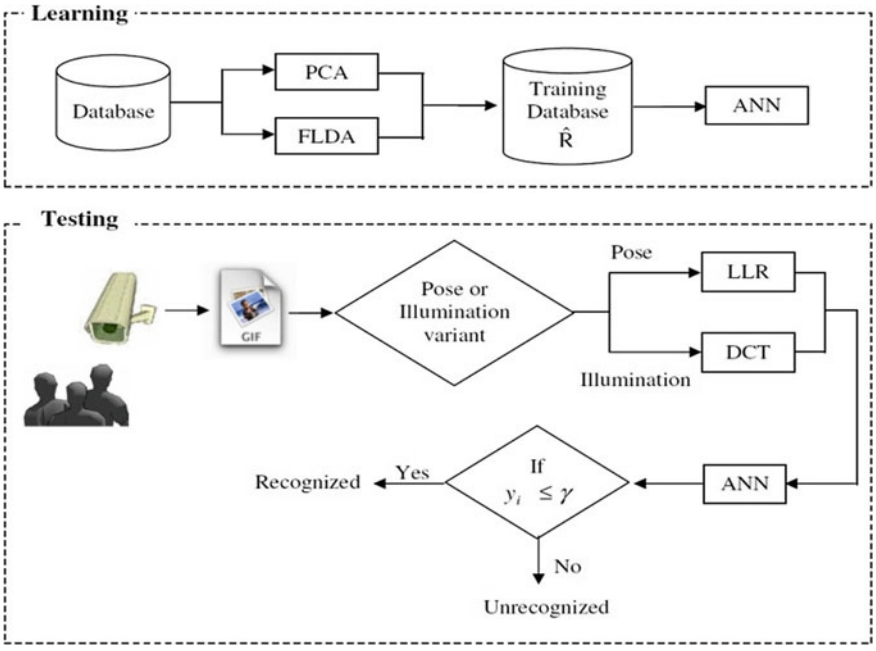


Fig. 8.2 LLR-based face recognition system

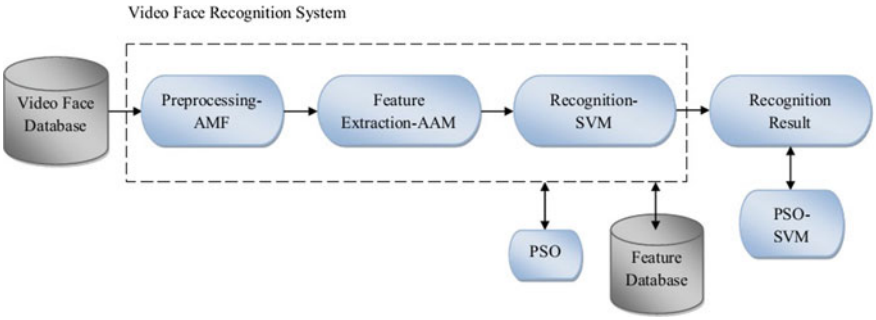


Fig. 8.3 Face recognition system based on PSO-SVM

8.5 Experiment and Result

In this section, the experimental outcome is described in detail, which is implemented in PC with 3.2 GHz Pentium Core i5 processor using MATLAB (version 7.2). To evaluate the effectiveness of the proposed method, the performance of proposed algorithm is evaluated on the UPC face database. This database has been made with the primary reason for testing the robustness of the face acknowledgment method against strong face posture variation and illumination variations. This database incorporates an aggregate of 44 people with 27 pictures for each individual which relate to various face posture views (0º, 30º, 45º, 60º, and 90º) under three distinct illumination conditions. Besides, this database incorporates 10 more extra frontal view of the face images with various occlusion and expression variations. The resolution of the pictures is 240 × 320 and they are in BMP format. Figure 8.4 shows the sample from UPC dataset.

This paper proposes LLR-based face recognition system [28] and PSO-SVM-based face recognition [29]. For evaluating the algorithms, we use UPC database. We divide the database into four datasets. The experiment is conducted to recognize the face images using all the four datasets.

From the analysis of results, it is clear that the proposed methods recognize the face images accurately. The proposed LLR-ANN method records an accuracy of 75%, and the second method PSO-SVM-based FR shows an accuracy of 95%.

Next, for comparison purpose, a recent AAM-based face recognition algorithm [30] is selected. This algorithm attains an accuracy of 90% in the experiment. The recognition performance of all the three methods is compared and tabulated in Table 8.1

Figure 8.5 shows the recognition performance of selected FR algorithms in terms of statistical measures.

As shown in Table 8.1 and Fig. 8.5, the experimental results indicate that PSO-SVM has higher face recognition performance compared to LLR-ANN-based method and AAM-based FR in terms of statistical measures. The performance



Fig. 8.4 Sample video frames from the UPC database

Table 8.1 Performance result of PSO-SVM and LLR method

Performance	PSO-SVM	AAM-based FR	LLR-ANN
Sensitivity	96.67	80.00	70.00
Specificity	93.33	100.00	80.00
Precision	93.55	100.00	77.78
NPV	96.55	83.00	72.73
FPR	06.67	0.00	20.00
FDR	06.45	0.00	22.22
FNR	03.33	20.00	30.00
Accuracy	95.00	90.00	75.00
F1-score	95.08	88.89	73.68
MCC	90.05	81.65	50.25

metrics show that the proposed PSO-SVM video face recognition technique is more accurately recognizes the face images than other two methods.

The computation time is also an important factor in face recognition. The time taken by LLR-based face recognition system [28], PSO-SVM-based methods [29], and AAM-based FR [30] are also compared, and the result is tabulated in Table 8.2. As can be seen from Table 8.2 that the PSO-SVM face recognition technique needs very less time for recognizing the faces compared to other two methods.

The main reason for selecting the UPC database is that it contains the face images of the individuals with different degrees of pose variation. The experiment is conducted on a dataset formed by selecting ten individuals. Table 8.3 tabulates the performance accuracy of all the three methods.

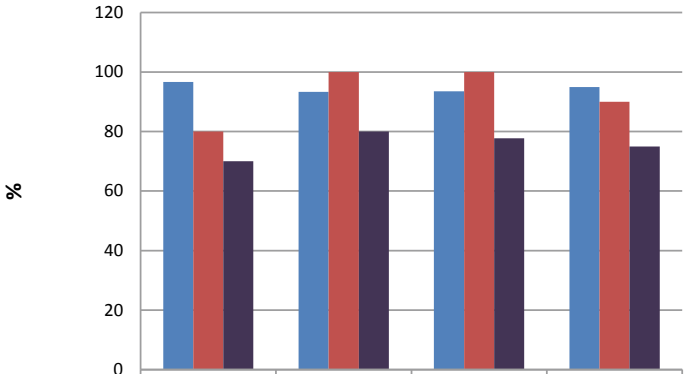


Fig. 8.5 Analysis of sensitivity, specificity, precision, and accuracy

Table 8.2 Comparison of computational time in terms of seconds

Experiment no	PSO-SVM	LLR-ANN approach	AAM-based approach
1	0.036011	08.1322	0.04679
2	0.047519	09.7772	0.05780
3	0.056591	10.5866	0.06501
4	0.054959	10.7222	0.06829
5	0.058237	10.1094	0.07184
6	0.049964	10.6153	0.06712
7	0.048922	09.4988	0.05793
8	0.050033	11.1353	0.07518
9	0.036495	08.6391	0.05745
10	0.055194	11.4592	0.07790

Table 8.3 Accuracy with different pose angle (%)

Pose angle/methods	0°	+30°	−30°	+45°	−45°	Avg
PSO-SVM	95.00	93.33	93.33	92.5	92.5	91.55
AAM-based FR	90.00	87.00	86.5	82.00	82.5	85.6
LLR-ANN	80.00	75.00	75.00	72.5	71.5	71.64

From Table 8.3, it is clear that when the pose variation is 0°, the recognition accuracy is high. The average accuracy of all the accuracy in each method is recorded separately for the comparison purpose. From Table 8.3, it is clear that the PSO-SVM performs better compared to other methods with an average accuracy of 91.55%.

The result analysis is given in Fig. 8.6. From Fig. 8.6, it is clear that PSO-SVM outperforms the LLR-based recognition system and AAM-based face acknowledgment by yielding a maximum accuracy of 95% and an average accuracy of 91.55%.

Table 8.4 shows the recognition performance of PSO-SVM and AAM-based methods in terms of statistical measures.

From Table 8.4, it is clear that PSO-SVM-based system has achieved a maximum accuracy of 95%, while AAM-based FR achieved 90% when the pose angle variation of 30 degree. For the pose variation of 45 degree and 60 degree, PSO-SVM-based

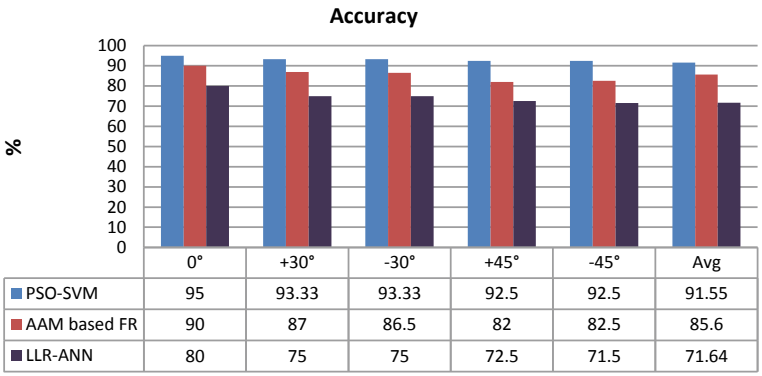


Fig. 8.6 Analysis of accuracy on different pose angles

Table 8.4 Statistical measures on different degrees of pose variations on UPC database

Statistical measures	PSO-SVM			AAM-based FR		
	30°	45°	60°	30°	45°	60°
Sensitivity	96.67	93.33	90.00	80.00	80.00	60.00
Specificity	90.00	90.00	80.00	100.00	80.00	80.00
Precision	96.67	96.55	81.82	100.00	80.00	75.00
NPV	90.00	81.82	88.89	83.00	80.00	66.67
FPR	10.00	10.00	20.00	0.00	20.00	20.00
FDR	03.33	03.45	18.18	0.00	20.00	25.00
FNR	03.33	06.67	10.00	20.00	20.00	40.00
Accuracy	95.00	92.50	85.00	90.00	80.00	70.00
F1-score	96.67	94.92	85.71	88.89	80.00	66.67
MCC	86.67	80.81	70.35	81.65	60.00	40.82

strategy shows an accuracy of 92.5% and 80%, respectively, which is comparatively higher rate than AAM-based method.

The above figures show the analysis of the proposed PSO-SVM-based FR [29] and AAM-based method [30] on different statistical measures. Figure 8.7 shows the accuracy, Fig. 8.8 shows recall, Fig. 8.9 shows specificity, and Fig. 8.10 shows recall. From the above figures, it is clear that the proposed PSO-SVM performs better compared to AAM-based face recognition.

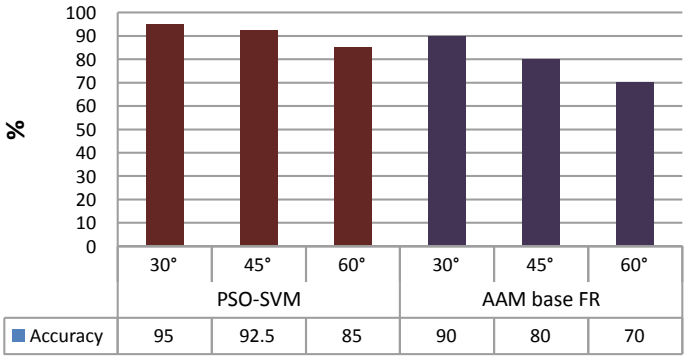


Fig. 8.7 Accuracy

Fig. 8.8 Sensitivity

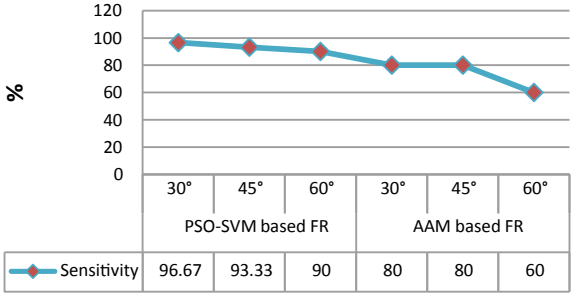


Fig. 8.9 Specificity

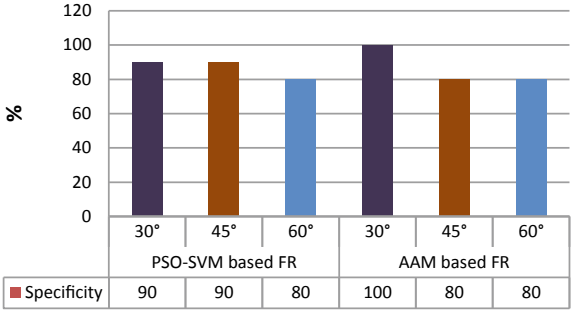
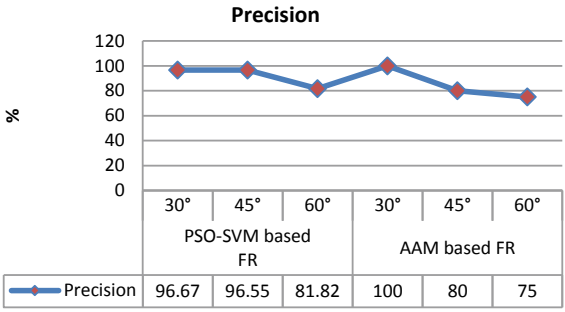


Fig. 8.10 Precision



8.6 Conclusion

This research work has proposed a basic face recognition technique to perform the face recognition task with the least process value. The work is expected to illuminate the disadvantage of high process value of the existing recognition techniques. The exploratory outcomes are incontestable on the grounds that the projected technique outperforms the past work concerning both, the recognition accuracy and the other computational complexities including the computation time.

In the integrative approach (LLR-ANN), initially, the PCA and LDA features of each image sample are combined and represented in feature space efficiently. During the testing phase, the image is tested for the illumination problem and pose variation. Illumination is normalized with DCT and pose is normalized with LLR. In any of these two cases, any way the feature vector is obtained using the combination of PCA and LDA. The test result of this method reported a maximum accuracy of 93.5%, an average recognition rate of 71.64%, and a recognition speed of 8.13 s.

In PSO-SVM-based approach**,** a new video face recognition system has been introduced. The proposed PSO is utilized to optimize the parameters of the SVM. In this method, the facial features are extracted from the face image using both AAM and shape model, and then the same is represented in feature space. These stored features are utilized to recognize the face image using SVM. This method yields a maximum accuracy of 95%, an average recognition rate of 91.55% with a minimum processing speed of 0.03second. By comparing the three selected methods, it is clear that the PSO-based face recognition stands first both in terms of statistical measures and recognition speed.

"We have taken permission from competent authorities to use the images/data as given in the paper. In case of any dispute in future, we shall be wholly responsible."

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. In: Appeared in IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics, vol. 14(1) (2004)
2. Delac, K.,Grgic, M.: A survey of biometric recognition. In: 46th International Symposium Electronics in Marine, ELMAR-2004, 16–18. Zadar, Croatia (2004)
3. Prabhakar, S., Pankanti, S., Jain, A.K.: Biometric recognition: security and privacy concerns. IEEE Sec. Privacy Magz. **1**(2), 33–42 (2003)
4. Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: Handbook of Fingerprint Recognition. Springer, NY (2003)
5. Jain, A.K., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society. Kluwer Academic Publishers (1999)
6. Turk, M., Pentland, A.: Eigenfaces for recognition. Cognitive Neurosci. **3**, 72–86 (1991)
7. Kirby, M., Sirovich, L.: Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Mach. Intell. **12**(1), 103–108 (1990)
8. Zhao, W., Chellappa, R.: Robust face recognition using symmetric shapefrom-shading. Technical Report, Center for Automation Research, University of Maryland (1999)
9. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: a literature survey. ACM Comput. Surv. (CSUR) **35**(4), 399–458 (2003)
10. Wu, Y.-L., Jiao, L., Wu, G., Chang, E.Y., Wang, Y.-F.: Invariant feature extraction and biased statistical inference for video surveillance. In: Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 284–289 (2003)
11. Ramanathan, N., Chellappa, R.: Face verification across age progression. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 462–469 (2005)
12. Chai, X., Shan, S., Chen, X., Gao, W.: Local linear regression (LLR) for pose invariant face recognition. EEE Trans. Image Process **16**(7), 1716–25. https://doi.org/10.1109/tip.2007.899195
13. Portera, G., Doran, G.: An anatomical and photographic technique for forensic facial identification. Forensic Sci. Int. **114**, 97–105 (2000)
14. Aggarwal, G., Roy-Chowdhury, A.K., Chellappa, R.: A system identification approach for video-based face recognition. Proc. Int. Conf. Pattern Recognit. **4**, 175–178 (2004)
15. Zhou, S., Krueger, V., Chellappa, R.: Probabilistic recognition of human faces from video. Comput. Vis. Image Underst. **91**, 214–245 (2003)
16. Stillman, S., Tanawongsuwan, R., Essa, I.: A system for tracking and recognizing multiple people with multiple cameras. In: Proceedings of International Conference on Audio and Video-Based Biometric Person Authentication, pp. 96–101 (1999)
17. Shermina, J.: Impact of locally linear regression and fisher linear discriminant analysis in pose invariant face recognition. Int. J. Comput. Sci. Netw. Sec. **10**(10) 106–110 (2010)
18. Zhang, H., Qu, Z., Yuan, L., Li, G.: A face recognition method based on LBP feature for CNN. In: 2017 IEEE 2nd advanced information technology, electronic and automation control conference (IAEAC), pp. 544–547 (2017)
19. Venkatesan, S., Srinivasa Rao Madane, S.: Face recognition system with genetic algorithm and ANT colony optimization. Int. J. Innov. Manage. Technol. **1**(5) (2010)
20. Xu, Y., Roy-Chowdhury, A., Patel, K.: Integrating illumination, motion, and shape models for robust face recognition in video. EURASIP J. Adv. Signal Process. p. 13 (2007)
21. Connolly, J.F., Granger, É., Sabourin, R.: An adaptive classification system for video-based face recognition, information sciences (In Press). https://doi.org/10.1016/j.ins.2010.02.026,march2010
22. Ng, H.U.: Pose-invariant face recognition security system. Asian J. Health Informat. Sci. **1**(1), 101–111 (2006)
23. Chai, X., Shan, S.: Locally linear regression for pose-invariantface recognition. IEEE Trans. Image Process. **16**(7), 1716–1725 (2007)

24. Arandjelovic, O., Cipolla, R.: A Methodology for Rapid Illumination-Invariant Face Recognition Using Image Processing Filters, pp. 159–171. Elsevier (2009)
25. Ding, C., Tao, D.: Trunk-branch ensemble convolutional neural networks for video-based face recognition. IEEE Trans. Pattern Anal. Mach. Intell. 1–14 (2017). https://doi.org/10.1109/TPAMI.2017.2700390
26. Wang, K., et al.: Facial Standardization Method in Video Face Recognition System (2016)
27. Liu, N.: Multiple Instance Learning with Deep Instance Selection for Video-based Face Recognition (2016). In: Dhiren, P., Jayesh, D. (eds.) PCA and LDA Method with Neural Network for Primary Diagnosis of Genetic Syndrome. International Advanced Research Journal in Science, Engineering and Technology, vol. 2, issue 10 (2015). https://doi.org/10.17148/IARJSET.2015.210. ISSN (Online) 2393–8021
28. Shreekumar, T., Karunakara, K.: Face pose and illumination normalization for unconstraint face recognition from direct interview videos. Int. J. Recent Technol. Eng. (TM), **7**(6S4) 59–68 (2019). ISSN: 2277- 3878 (Online)
29. Shreekumar, T., Karunakara, K.: A video face recognition system with aid of support vector machine and particle swarm optimization (PSO-SVM). J. Adv. Res. Dyn. Control Syst (JARDCS) **10**, 496–507 (2018)
30. Prasanna, K.M., Rai, C.S.: A new approach for face recognition from video sequence. In: 2018 2nd International Conference on Inventive Systems and Control (ICISC), pp. 89–95. Coimbatore (2018)
31. Shreekumar, T., Karunakara, K.: Active appearance model based for face recognition from surveillance video. Test Eng. Manage. **83**, 6969–6981 (2020). ISSN:0193–410
32. Shreekumar, T.,Karunakara, K.: Face pose and blur normalization for unconstraint face recognition from video/still images. Int. J. Innov. Comput. Appl. Indersci (2020). ISSN:1751–648X
33. Shreekumar, T., Karunakara, K.: Identifying the faces from poor quality image/video. Int. J. Innov. Technol. Exploring Eng. (IJITEE). **8**(12), 1346–1353 (2019). ISSN: 2278–3075.
34. Amrutha, K., Shreekumar, T.: Instant warning system to detect drivers in fatigue. Int. J. Sci. Res. (IJSR), **4**(2), 791–79 (2015)

Chapter 9
Application of Bat Optimization Algorithm for Power System Loss Reduction Using DSSSC



M. Thirumalai and T. Yuvaraj

Abstract A novel method is proposed for finding the optimal location along with the sizing of DSSSC in order to mitigate the power losses of the system. DSSSC sitting and sizing are determined by bat algorithm. Testing of the pro-posed work is done on the IEEE 69-bus which is standard system of radial bus distribution. The results obtained shows that the placement of optimal point of DSSSC with its sizing has reduced the power losses of the total radial distribution network system.

9.1 Introduction

The power system consist of three sections namely generation of power, transmission of power and distribution of power, among which resistance to reactance ratio is higher in the case of distribution which results in instability of the voltage and more power loss. The conventional methods for the maintenance of acceptable voltage stabilization in the distribution system use series voltage regulator and shunt capacitor [1–3]. But there are some demerits with these conventional methods as reactive power generation is not possible by the series voltage regulator and operates step by step leading to slower response [4]. Also variable reactive power is not generated by the shunt capacitors continuously. The behavior of the distribution capacitors are oscillatory when used along with inductive components within the same circuit is also a problem to be focused. Introduction of devices based on the concept of FACTS resolves the defects due to series voltage regulators and the shunted capacitors. The FACTS concept was initially introduced for transmission systems but later applied also for distribution systems to mitigate the system losses, power factor correction and for improved voltage of the power system.

M. Thirumalai · T. Yuvaraj (✉)
Department of Electrical and Electronics Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

M. Thirumalai
Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_9

A synchronous generator of static type namely static synchronous series compensator (SSSC) operates without electrical energy as a controllable voltage is injected by the series compensator using a VSC where the inoculated voltage is in quadrature with the power network's route current. Hence rapid impedance compensation is provided both capacitive and inductive by the device without depending on the line current of the power network. DSSSC is comparatively better than conventional compensation devices for reactive power as it has various features namely modification of power factor by inoculation of voltage continuously together with organized regulator, balancing of load in distributed organized systems, covering the power demand both reactive & capacitive, control of power flow, reduced harmonic distortion by the process of active filtering [5, 6]. Designing SSSC based controller based on the technique of GA for improving the transient performance using the approach of multi-objective optimization in power system is done in [7]. Sizing of SSSC controller is done by optimized method in the network of transmission [8].

Reduction of transmission losses is done by the PSO technique which modifies the algorithm of Newton—Raphson load flow improving the stability using SSSC damped controller for minimizing the objection function based on the time domain. Oscillation damping to power system is done efficiently to develop the bus voltage [9]. In the system of radial distribution, locating the optimum point of DSTATCOM has greater influence. If the placement is done improperly then the benefits can be reduced which will still endanger the operation of the entire system [10].

So far only one researcher used DSSSC in distribution system for loss mitigation. The approach of particle swarm algorithm is implemented in [12] which identify the appropriate allocation of DSSSC with reduced power loss as objective function was investigated.

The proposed method implements bat optimization algorithm for identifying the best location of DSSSC along with its sizing in order to mitigate the total system losses in the RDS. Testing of the present method is done on IEEE 69-bus on RDS. Evaluation of the outcomes acquired from the present approach with that of the already available approaches is done and found to be better.

9.2 Formulation of the Problem

9.2.1 Analysis of Load Flow

The conventional methods for studying the load flow such as GS method, NR technique and the method of Decoupled load flow are not suited for obtaining the line flow & voltages in the system of RDS as the ratio of R/X is higher. A straight methodology for the power flow for the distribution system is discussed in [13]. A simple distribution system is represented by the single line diagram (see Fig. 9.1).

At node t, the current injected is obtained from Fig. 9.1 as follows

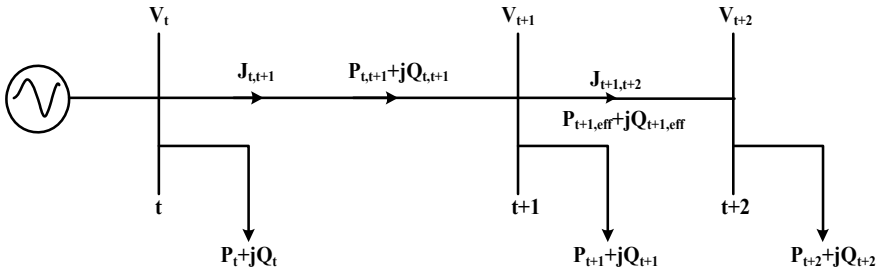


Fig. 9.1 One line diagram

$$I_t = \left(\frac{P_t + jQ_t}{V_t}\right)^* \tag{9.1}$$

The branch current is determined using KCL between the nodes t & $t+1$ from Fig. 9.1 which is obtained as,

$$J_{t,t+1} = (I_{t+1}) + (I_{t+2}) \tag{9.2}$$

By utilizing the Bus Injected to Branch Current matrix, the Eq. (9.2) is signified in matrix form as below.

$$[J] = [I]\,[\text{BIBC}] \tag{9.3}$$

The bus voltage at $t+1$ is determined using KVL from the Fig. 9.1 which is obtained as,

$$V_{t+1} = V_t - \left[J_{t,t+1}(R_{t,t+1} + jX_{t,t+1})\right] \tag{9.4}$$

The real power loss and imaginary power loss between the nodes t & $t+1$ in the line section can be obtained as

$$P_{\text{Loss}(t,t+1)} = \left(\frac{P_{t,t+1}^2 + Q_{t,t+1}^2}{\left|V_{t,t+1}\right|^2}\right) * R_{t,t+1} \tag{9.5}$$

$$Q_{\text{Loss}(t,t+1)} = \left(\frac{P_{t,t+1}^2 + Q_{t,t+1}^2}{\left|V_{t,t+1}\right|^2}\right) * X_{t,t+1} \tag{9.6}$$

In the RDS, the average system loss is obtained by adding the losses occur-ring in all the line section as follows,

$$P_{T\,\text{Loss}} = \sum_{t=1}^{b} P_{(\text{Loss},t,t+1)} \tag{9.7}$$

9.2.2 Power Loss Mitigation

The aim of allocation of DSSSC in the RDS is to mitigate the losses while satisfies the constraints of equality and in equality. The main aim of the work is given by

$$\text{Minimize}(F) = \text{Min}(P_{T\,\text{Loss}}) \tag{9.8}$$

9.3 Proposed Bat Algorithm

9.3.1 Overview of Bat Algorithm

The most power algorithm for the optimization problems of tough power system are the nature inspired algorithms proposed by Yang [11]. Bats are the only mammals with wings have the special feature of finding their prey with advanced capability of echo location. The objects in the surrounding are detected by the radiation of the sound signal known as echo location so that way is found even when it is completely dark.

Development of Bat algorithm is idealized by the appearances of bats. The three rules romanticized are assumed as:

1. The characteristic of echo location is utilized by each bat to sense the distance with the knowledge of differentiating prey and other obstacles by nature.
2. The motion of bat is random with velocity u_i, position y_i, frequency $g_{\min}$, variable wavelength λ & loudness B_0 for seeking the prey. Bats have the ability of regulating the frequency of the emitted pulse also to regulate the pulse emission rate r within the array of [0, 1].
3. Though there is possibility for the variation in loudness, assumption is done that the loudness differs from positive great value B_0 to a constant least value $B_{\min}$.

9.3.2 Population Initialization

Initially for the BA population is produced arbitrarily hence the no of virtual bats are randomly selected. The population of the simulated bats can be between 10 & 40. Once after the fitness done initially, updating of the values of pulse rate, loudness, movement is performed for a given function.

9.3.3 Motion of Virtual Bats

The rules are defined in bat algorithm for updating the velocity u_i and position y_i of the simulated bats which are assumed as follows:

$$g_i = g_{\min} + (g_{\max} - g_{\min})\beta \tag{9.9}$$

$$u_i^t = u_i^{t-1} + (y_i^{t-1} - y^*)g_i \tag{9.10}$$

$$y_i^t = y_i^{t-1} + y_i^t \tag{9.11}$$

where the random vector $\beta \in [0, 1]$ obtained from uniform distribution where y^* is the best location indicating the current solution out of the "n bats".

The innovative result generated locally for all the bats with arbitrary walk is specified by the Eq. (9.12).

$$y_{\text{new}} = y_{\text{old}} + \varepsilon A^t \tag{9.12}$$

where ε is the arbitrary no in the assortment [0, 1], while $B^t = (B_i^t)$ is the middling loudness of all the bats at this time step.

9.3.4 Emission of Pulse and Loudness

The loudness B_i is proceeded based on the iteration with pulse emission rate r_i is updated when a bat is nearing its prey, the volume reductions with increase in the rate of pulse production hence the convergence equation is given as

$$B_i^{t+1} = \alpha B_i^t \tag{9.13}$$

$$R_i^{t+1} = r_i^0[1 - \exp(-\gamma t)] \tag{9.14}$$

where α and γ are continual values.

For any instant of "$0 < \alpha < 1$" and $\gamma > 0$ we have.

$B_i^t \to 0, r_i^t \to r_i^0$ as $t \to \infty$.

The starting value of loudness B_0 can characteristically be in the sort of [0, 1], while the starting value of emission rate r_i can be in the sort of [0, 1].

Bat Algorithm is implemented step by step based on the approximation discussed above as follows:

Step 1: The input data namely the bus data and load data are read by the system.

Step 2: The load flow distribution of the base case is run to determine to the reactive power loss, real power loss along with the voltages.
Step 3: The applicant buses are identified for the location of DSSSC.
Step 4: The control parameters of the bat algorithm such as loudness, pulse rate, pulse frequency, maximum number of iterations are set along with the inferior and higher bound of the constraints.
Step 5: The population of bat initially is generated in a random manner. The optimal size of each bat is encouraging (p.u) for the devices of DSSC in the distributed network.
Step 6: The suitability function is evaluated. Calculation of the expected reactive power losses, active losses and the objective function voltage deviance is done for each bat or solution using the method of direct load flow.
Step 7: The bat with minimum value of power loss is selected as the best bat among the population.
Step 8: The population of bat is updated.
Step 9: The load flow is run for determining the unknown true and imaginary power loss values.
Step 10: The criteria for closure is checked which can be extreme no of repetitions for updating the populace of the BA or the value for which the OF will be least. Once the condition is satisfied the algorithm ends else returned to step 5.
Step 11: Optimal solutions are displayed.

The above steps in the bat algorithm are followed for minimizing the objective function.

9.4 Discussion on Results

9.4.1 Test System for IEEE 69-Bus

The present method shows performance proving its effectiveness by testing on standard IEEE 69-bus system. The test system involves of 69 buses & 68 divisions hence a large scale system with a total reactive load of 2.69 MVAr and total real load of 3.80 MW respectively. The IEEE 69-bus radial distributed system is displayed with its one line illustration in Fig. 9.2. The proposed method is simulated using MATLAB which runs the MATLAB code of the RDS to calculate the reactive power loss and the real power loss for identifying the optimum sitting of DSSSC and the sizing of DSSSC. In this work, the optimal installation of DSSSC is done by selecting the 61st bus system.

The comparison of the locations, real power losses, voltage values of the existing methods (p.u) & the present approach is done in Table 9.1. The proposed method has reduced the real power loss to 24 kW which indicates 89.39% of mitigation of power loss after the installation of DSSSC. This mitigation in power loss is high compared to that of the existing methods. The IEEE 69-bus bus voltage diagram with DSSSC

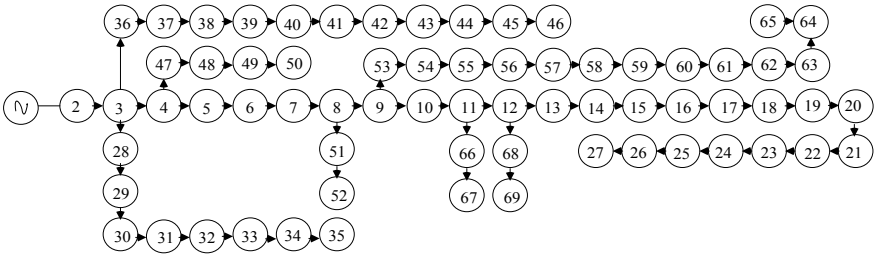


Fig. 9.2 One line diagram of IEEE 69-bus system

Table 9.1 The results of DSSSC allocation for a 69-bus system

	Uncompensated	Analytical method	PSO[12]	BA
KVAr (p.u)	–	1.9229	2.1377	2.0737
Location	–	60	60	61
Power loss (kW)	225.4	125.5	37	24
% of power loss reduction	–	44.32	83.58	89.35
Voltage (p.u)	0.9196	1.0	1.0065	1.0

& without DSSSC is shown in Fig. 9.3. Effective improvement is viewed in the bus voltage diagram of the RDS as implemented with DSSSC. Hence optimization based on the present BA approach is efficient related to the prevailing approaches.

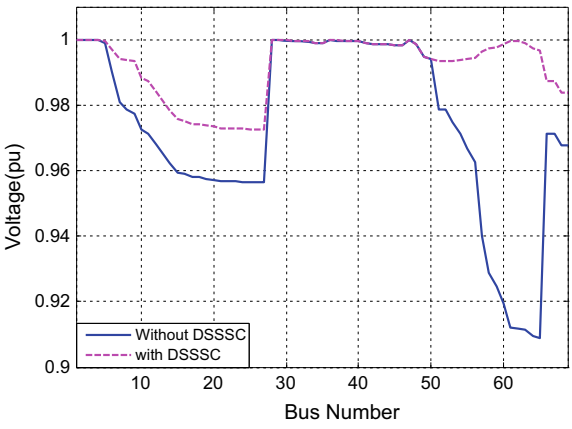


Fig. 9.3 Bus voltage diagram improvement for 69 bus system

9.5 Conclusion

A new BA based approach is presented in this paper which finds the optimal location of DSSSC along with its sizing of the RDS. Testing of the present approach is done on standard IEEE 69—bus system. Proper location and sizing of the DSSSC by present approach ensures the greater loss mitigation. The simulated results show that the implementation of the proposed method with DSSSC has reduced the total real and imaginary power losses better than the existing methods. Bat algorithm is accurate in finding the solution for optimal location which can be applied for *n* no of buses.

References

1. Casillas, G.I., et al.: Voltage regulators, capacitor banks and distributed resources allocation in a distribution network system. In: 2017 IEEE PES Innovative Smart Grid Technologies Conference-Latin America (ISGT Latin America). IEEE (2017)
2. Sarfi, V., Livani, H.: Optimal Volt/VAR control in distribution systems with prosumer DERs. Electr. Power Syst. Res. **188**, 106520 (2020)
3. Xie, Q., et al.: Use of demand response for voltage regulation in power distribution systems with flexible resources. IET Generat. Trans. Distribut. **14**(5), 883–892 (2020)
4. Yuvaraj, T., Devabalaji, K.R., Thanikanti, S.B.: Simultaneous allocation of DG and DSTATCOM using whale optimization algorithm. Iranian J. Sci. Technol. Trans. Electri. Eng. **44**(2), 879–896 (2020)
5. Hingorani, N.G., Gyugyi, L.: Understanding FACTS; Concepts and Technology of Flexible AC Transmission Systems. IEEE Press book (2000)
6. El-Zonkoly, A.: Optimal sizing of SSSC controllers to minimize transmission loss and a novel model of SSSC to study transient response. Electr. Power Syst. Res. **78**(11), 1856–1864 (2008)
7. Sidhartha, P.: Multi-objective evolutionary algorithm for SSSC-based controller design. Electr. Power Syst. Res. **79**, 937–944 (2009)
8. El-Zonkoly, A.: Optimal sizing of SSSC controllers to minimize transmission loss and a novel model of SSSC to study transient response. Electr. Power Syst. Res. **78**, 1856–64 (2008)
9. Panda, S., Padhy, N.P., Patel, R.N.: Power-system stability improvement by PSO optmized SSSC-based damping controller. Electr. Power Compon. Syst. **36**, 468–90 (2008)
10. Yuvaraj, T., Ravi, K., Devabalaji, K.R.: DSTATCOM allocation in distribution networks considering load variations using bat algorithm. Accepted on 26 September 2015. Ain Shams Eng. J. Elsevier
11. Yang, X.-S.: A new meta heuristic bat-Inspired algorithm. In: Gonzalez, J.R. et al. (eds.) Nature Inspired Cooperative Strategies for Optimization (NISCO 2010). Studies in Computational Intelligence, pp. 65–74. Springer, Berlin (2010)
12. Devi, S., Geethanjali, M.: Optimal location and sizing of distribution static synchr nous series compensator using particle swarm optimization. Electri. Power Energy Syst **62**, 646–653 (2014)
13. Jen-Hao, T.: A direct approach for distribution system load flow solutions. IEEE Trans. Power Deliv. **18**(3), 882–887 (2003)

Chapter 10
Malaria Cell Detection Model



Prasannajeet Bajpai and Sanchita Chourawar

Abstract Malaria is the most ubiquitous endemic and a deadly disease which is caused by parasites which are transmitted to people through bites of an infected female Anopheles mosquito. This disease is present in several continents, which is a matter of concern for several bodies and associations. All the governments around the world are trying to reduce this morbidity by taking various preventive measures. Model will help us to detect five species of Plasmodium (Malaria Parasites) that have long been recognized to infect humans. Now with our model which is a pre-trained CNN model with 13,775 images of infected cells and 13,775 images of uninfected cells with valid efficiency more than 93% and then all it is needed is just upload image or images that we want to test which then will be fed to our mode and then the model will give the output with concerning images in the same sequence in which it was uploaded. By using this model, we are primarily promoting reusability, accessibility, cost-efficient, eco-friendly as we are not using any kits to create more plastic waste.

10.1 Introduction

Malaria is caused by Plasmodium parasites. These parasites spread to people by the bite of infected female Anopheles mosquitoes also called to as "malaria vector." Five parasites species which are found in humans, and a couple of those species are—*P. falciparum* and *P. vivax* which pose a great threat. With over quite 200 million cases per annum and most of its cases are from remote or rural areas of sub-Saharan Africa South-East Asia, Eastern Mediterranean, Western Pacific, and therefore, the Americas where immediate testing is not always possible but by an invention of 20% "hand-powered blood centrifuge" by Manu Prakash which can reach up to 125,000 rpm enough to separate plasma from blood after which we can directly capture an image of an individual cell and then feed it to our model after which the model will give us the result. There are only three essential steps for the implementation of this model:

P. Bajpai · S. Chourawar (✉)
Government Engineering College Bilaspur, Bilaspur, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_10

(a) Collect and Arrange Data:

- First, we will gather multiple images of parasitic and uninfected cells.
- Now we classify them into two categories "train" and "test."
- Then inside both test and train, we will keep parasitic and uninfected cells images which will be used to rain our model properly.

(b) Training the Model:

- Before doing anything, we will need to import all important libraries like shutil, TensorFlow, matplotlib.
- Now after having all the libraries imported, we will have to make our model which will be "keras.models.Sequential" in this case.
- And then we will optimize it with RMSprop as an optimizer.
- Then the final step is to fit this model which gives use of valid-accuracy more than 94.

(c) Using the Model:

- Now, all we need to do is just upload the image.
- Then we will get results corresponding to our images in the same order in which we uploaded.

10.2 Literature Review

ImageNet classification which is a set of a large number of images and with the help of deep convolutional neural network this is where we can say that it all started [1]. Albeit deep learning has been around since the 70s with AI heavyweights Geoff Hinton, Yann LeCun, and Yoshua Bengio giving all their time and efforts on convolutional neural network until AlexNet (2012) brought deep learning into the mainstream where he builds an easy model to differentiate between cat, dog, and duck [2, 3]. Usage and hindrance of malaria rapid diagnostic testing (2009) within the war-torn The Democratic Republic of Congo made a search paper where they concluded that CHWs will be allowed to use RDTs for the management of febrile children safely and effectively; but, the cost-efficiency of RDTs is restricted in the area of high malaria prevalence [4, 5]. Multimodal Semantic Division: The fusion of RGB and Depth Data in CNN research by Zoltan Kopanyi, Dorota Iwaszczuuk, Bing Zhao, Can JozefSaul, Charles K. Toth, Alper Yilmaz (2019) where they showed how an RGB image goes through a number of phases it goes through where it shows Feature Extraction, Thresholding, Edge Detection, and many more which are important for classification of an image [6]. From the research paper to deep learning model in Keras by using python for image segmentation by Peterfitchcsiro (2020) where we use convolutional neural network to take input and then pass it through multiple layers of Conv2D and MaxPooling2D to create a proper sequential model [7, 8]. Malaria: The Past and the Present by Jasminka, Talapko, Ivana Skrlec, Tamara Alembić, Melita,

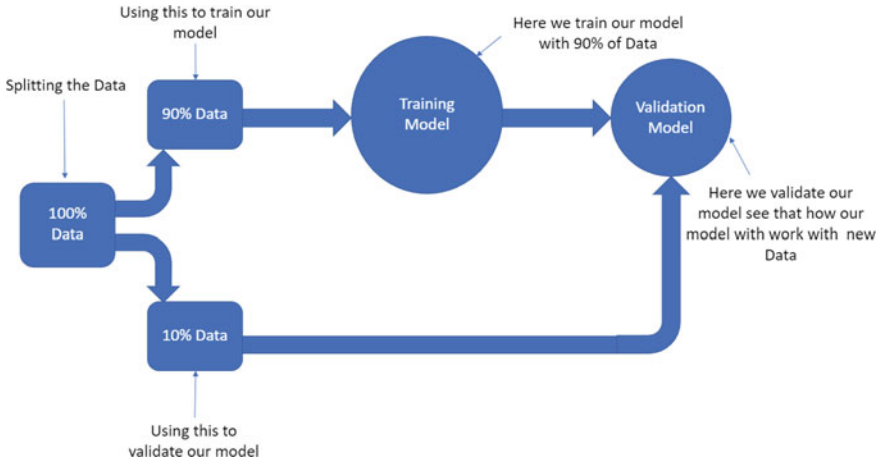


Fig. 10.1 Flowchart of "Training and Validating Model"

Jukić, Aleksandar, Vćev by (2019) which escalated our knowledge about this disease [9]. Having an alternative solution like using a pre-trained CNN model will help us overcome many odds like reducing the use of plastic which is a promising step to a future generation, providing easy access to testing as we need only a microscope and a camera to capture the image of a cell and the camera could also, be of a smartphone which would make it easier for anyone to access it an as large number of the population has access to smartphones and the government also making huge efforts to make them available to everyone, with such a high efficiency we can easily rely on models like this and in near future, we can gather more data and then we won't just be using it to detect malaria in fact when can use it detect multiple diseases [10] (Fig. 10.1).

10.3 Architecture and Working

Now moving any further, we want to show how we trained our data and how exactly it works initially we collected a set of data where we have thousands of images of parasitic cells and uninfected cells. Now we use a simple split function where we will be splitting the data into two parts where one will be a training dataset and another one will be for testing or validating our model which gives a proper picture of how our model will perform with new data.

Further when it comes to the model building which here in the case will be a "Keras.models" and hyperparameter is not given meticulous attention in most of the frameworks in machine learning, but intuitively they govern the underlying system on a "High Level" than the primary parameters of interest and our CNN model consists of 3 Conv2D Layers, 3 2DMaxPooling Layers, 2 Dense Layers, and a Flatten Layer.

Conv2Dlayer produces a convolution kernel that's involved in the layer input to build a tensor of outputs. If use bias is True, a learning vector is formed and is attached to the outputs. The primary requirement within the Conv2D parameter is the sum of filters from which the convolutional layer learns. Layers early within the specification learn lesser from convolutional filters while from the layers deeper within the network (i.e., near to the output predictions), it will study better with filters. Conv2D layers in middle of this will grasp more information from this filter than the first Conv2D layers but fewer filters than the layers closer to the output that's why we would like to use multiple layers during a model for correct working.

2D MaxPooling reduces the dimensions of the data, the quantity of parameters, the quantity of computation needed, and it also controls overfitting. You'll remark that 2D MaxPooling is analogous to lower the dimension of an image. This layer outputs a smaller tensor than its input, which suggests downstream layers will need fewer parameters and amount of computation; it also serves to regulate overfitting. It has five parameters which are used to:

- Horizontal Pooling Factor
- Vertical Pooling Factor
- Horizontal Stride
- Vertical Stride
- Padding.

Flatten Layer is used to flatten the input as we applied for an input batch size of (2,2) which will be converted to a batch size of 4 what a flatten layer is that it converts our CNN to DNN for efficient training (Fig. 10.2).

Dense layer is one of the most frequently used layers and is also known as "regular deeply connected neural network layer."

The optimizer is one of the most important hyperparameters for tweaking weight of the model and gets pretty optimized results for our model. We have a variety of optimizers for optimizing but here we have used RMSprop with lr = 0.001 but in options, we also have Adagrad, SGD, Nesterov Momentum, Adam, BFGS, etc.

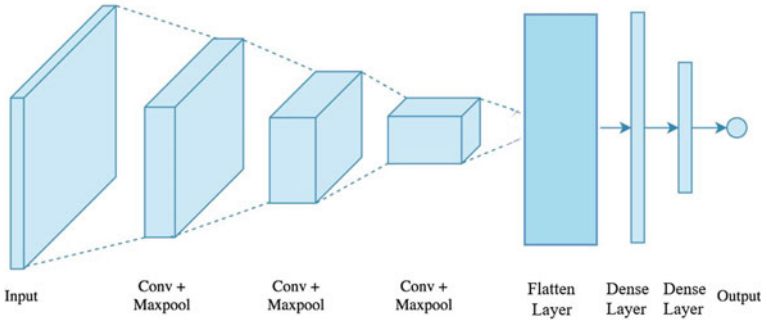


Fig. 10.2 Flowchart of sequential model

```
[ ]  history = model.fit(train_generator,
                         validation_data=validation_generator,
                         steps_per_epoch=100,
                         epochs=5,
                         validation_steps=50,
                         verbose=2)
```

```
Epoch 1/5
100/100 - 8s - loss: 0.7438 - accuracy: 0.5365 - val_loss: 0.6422 - val_accuracy: 0.6460
Epoch 2/5
100/100 - 8s - loss: 0.6597 - accuracy: 0.6070 - val_loss: 0.6695 - val_accuracy: 0.5550
Epoch 3/5
100/100 - 8s - loss: 0.5551 - accuracy: 0.7285 - val_loss: 0.3889 - val_accuracy: 0.8470
Epoch 4/5
100/100 - 9s - loss: 0.3342 - accuracy: 0.8750 - val_loss: 0.2639 - val_accuracy: 0.8980
Epoch 5/5
100/100 - 8s - loss: 0.2487 - accuracy: 0.9200 - val_loss: 0.2039 - val_accuracy: 0.9340
```

Fig. 10.3 Fitting the model



Fig. 10.4 Graphical representation of "Training and Validation Accuracy" throughout epoch cycles

Then we will have to fit the model in which steps per epoch is been set as 100 with epochs as 5 and 50 validation steps with 2 verbose, after 5 epochs cycles, we got a model with + 94% in its validation (Fig. 10.3).

Here to accomplish such a high accuracy at such a less time is possible due to a high GPU utilization and as we build this model on an online platform so we were gifted with a high-end GPU card "GTX TITAN X" which has more than 3000 cores which is a lot to work with these number of data (Figs. 10.4, 10.5, 10.6 and 10.7).



Fig. 10.5 Graphical representation of "Training and Validation Loss" throughout epoch cycles



Fig. 10.6 Snap of cell images

Choose Files 4 files

- **C4.png**(image/png) - 14451 bytes, last modified: 01/07/2020 - 100% done
- **C3.png**(image/png) - 13727 bytes, last modified: 01/07/2020 - 100% done
- **C2.png**(image/png) - 8568 bytes, last modified: 01/07/2020 - 100% done
- **C1.png**(image/png) - 8536 bytes, last modified: 01/07/2020 - 100% done

```
Saving C4.png to C4 (1).png
Saving C3.png to C3 (1).png
Saving C2.png to C2 (1).png
Saving C1.png to C1 (1).png
[0.]
C4.png is a parasitic
[0.]
C3.png is a parasitic
[1.]
C2.png is uninfected
[1.]
C1.png is uninfected
```

Fig. 10.7 Output of our model after prediction

10.4 Result Analysis

Now we come to the final step where we will put some new untrained images.

Now after successfully uploading images, our model will use its pre-trained data and will display us the output as shown below.

Now deploying this model is very easy all we have to do is just convert it into either .h5 or tflite model. Tflite is always a better option if we are trying to make a model which is to be made pretty portable because of the size difference between .h5 model and tflite model, and tflite model be implemented in android applications also make our model more effortless and convenient for use.

10.5 Conclusion

By implementation of this model, we are able to make a new tech for the detection of malaria disease and also an alternative for the kits which are traditionally used. Our model has lots of merits because with our we are be able to achieve more accurate results as some of the kits are faulty or damaged. It is also reusable and eco-friendly as we do not have to litter kit's waste. By implementing our model in an android device or any other micro-controller, we can increase its compatibility and mobility. Our model is just one-time investment as we do not have to buy any equipment. It is also expandable because currently, we are working with just malaria cells but in future we might be able to detect multiple diseases making our model more versatile with some more dataset.

Acknowledgements In the present world of completion, there is a contest of living in which those that have the desire to come forward will only succeed, and with a model like this, we construct a bridge between theoretical and practical working. Like there is a saying "you never know your limits unless you push them," and this project made me realize my limits about which I am pretty sure that I will be able to push it further later. A teacher is a compass that activates the magnets of curiosity, knowledge and wisdom in their students and I am so fortunate that I have Professor Sanchita Chourawar (Assistance Professor, Department of Computer Science and Engineering). A special thanks to our Head of Department Professor Sourabh Yadav (Department of Computer Science and Engineering) for his constant cheer up and wholehearted support. I am deeply indebted and that I would really like to manifest my sincere gratitude to our honored Head of Institute Dr. B. S. Chawla (Principal, Government Engineering College, Bilaspur) for providing me with an opportunity to do this project.

References

1. AlexNet.: ImageNet classification with the help of CNN **115**, 11–97 (2012)
2. Community health employee of democratic republic of congo for usage and hindrance of malaria rapid diagnostic testing. Malaria J. **8**(308) (2019)
3. Hinton, G., LeCun, Y., Bengio, Y.: Deep learning. Nature **521**(7553), 436–44 (2019)

4. Tseng, C., Lee, S.: A supervised learning method for the design of linear phase FIR digital filter using Keras, pp. 1–2 (2019)
5. Talapko, J., Skrlec, I., Alebić, T., Jukic, M., Včev, A.: Malaria: the past and the present **7**(6):179 (2019)
6. Hinton, G.E., Krizhevsaky, A., Srivastava, N., Sutskever, I., Salakhuttdinov, P.R.: Dropout an easy method to avoid neural networks from overfitting **15**(56), 1929–1958 (2014)
7. Peer, J.: Image processing in python **19**(2), e453 (2014)
8. Jai, X.: Image recognition method based on deep learning. In: CCDC (2017)
9. Lina, B.: Deep Learning with Keras (2017)
10. Gajurel, A., Louis, S.J.: GPU acceleration in sparse neural networks (2020). arXiv:2005.04347 [cs.DC]

Chapter 11
Design and Analysis of Fractal Monopole Antennas for Multiband Wireless Applications



Siddhi Oja and P. M. Menghal

Abstract A design of fractal monopole patch antenna using metamaterial-based substrate and Minkowski fractal has been presented in this paper. The antenna has been designed for a frequency range from 1 to 6 GHz which is mainly covering the wireless fidelity and Wi-Max ranges. Antenna parameters like VSWR, return loss, and bandwidth have been analyzed for the proposed antenna. The overall results of proposed antenna in design after carrying out parametric study are found: gain above 4 dB, return loss (S−11) −37.11 dB ($f_L = 0.3$) and −59 dB ($f_L = 0.2$), VSWR −1.02, and bandwidth −77.01%. The simulated characteristics of the antenna along with the 3D radiation patterns and gain are presented and discussed by using CST simulation software**.**

11.1 Introduction

An antenna is an indispensable device for any wireless communication system. Swift development in satellite and wireless communication mainly in commercial as well as government communication systems in the last few years has compelled more and more requirement of developing low value, low profile, minimal weight, and broadband antennas that are capable of maintaining high performance over an extensive range of frequencies. The most common version of a printed antenna is a microstrip antenna or patch antenna. In this type of antenna, a metallic patch is designed in coordination with ground and its dimensions are appropriately chosen as per the requirement of operability in the particular frequency domain. These antennae are more promoted because of their prostrate cost, simple fabrication, and effortless integration with circuit components [1]. The most exclusive attribute of these antennas is that they are conformal to a planer and non-planer surfaces, hence can be easily mounted on curvy surfaces as of aircraft, missiles, space crafts, satellite systems

S. Oja · P. M. Menghal (✉)
Faculty of Electronics, Military College of Electronics and Mechanical Engineering, Secunderabad 500015, Telangana, India
e-mail: prashant_menghal@ieee.org

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_11

radars, and mobile devices [2–4]. Also, it is highly versatile in terms of polarization, pattern, and impedance, once an operable mode of patch and form is chosen as per various requirements. The versatile designs that can be achieved with microstrip antenna in all probability exceed that of the other kind of antenna elements [4–8]. In this paper, a flexible monopole fractal antenna is designed using flexible Rogers RO4003C as a substrate. Minkowski fractal geometry is used as a base, which is modified to design the new geometry of the antenna.

11.2 Fractal Monopole Antennas

Fractal is a technology that can lead to antenna miniaturization, multifrequency, and ultra-wideband applications, and this paper is talking about the same. The existing communication system is evolved to broadband and integration, but the people's need for portable mobile communication has increased. This needs antenna development depending upon the broadband technology, multifrequency, and miniaturization technology. Multifrequency antenna is gauging attention, owing to its small volume, lighter in weight, easy, and active circuitry integration advantages, especially due to the emerging wireless communications [2–4]. For antenna operating on different frequencies, with small size, operable in wideband and has a circular polarization, such technology will be the domestic and foreign research key areas [5, 6]. In times to come, communications will lead to some serious technical challenges in the field of antenna design, demanding a miniaturized structure to execute its tasks. The fractal antenna can cater to the need for antenna requirements necessary for modern communication devices with a very thin section, small size, and is easy to manufacture and cheap price [7, 8]. Fractal is produced with the self-similarity of fractal dimension structure through the number of iterations. It will not only lead to reducing the dimensions of the antenna, but also strengthened the characteristics of a directional antenna. Minkowski fractals are given by Hermann Minkowski in the form of representation and definition of geometries in the year 1885 [9, 10]. The square patch geometry on being further compressed into five smaller squares leads to the formation of the Minkowski fractal geometry given in Fig. 11.1. This is explained in terms of affine transformation here.

11.3 Antena Design

The design has been evolved after a series of steps starting from designing a simple patch antenna to a monopole patch antenna, followed by applying Minkowski fractal and finally modifying the Minkowski geometry to achieve the final design. The antenna parameters taken into consideration to study the results of the various designs are return loss, VSWR, and gain. The antenna design equation is given in Appendix

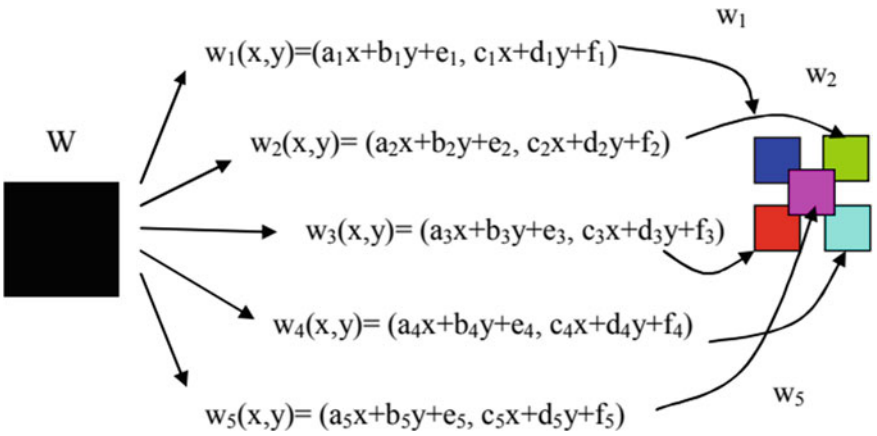


Fig. 11.1 Generation of Minkowski fractal geometry

'A' The design specification of the proposed antenna based on the calculation as mentioned above is given in Table 11.1.

The process of design evaluation of antenna is shown in Fig. 11.2.

11.4 Simulation Results and Comparative Analysis

In the proposed antenna design, different substrate techniques and multilayer substrates are used to enhance the bandwidth, return loss (S−11), and other properties of the antenna. The proposed multilayer monopole patch antenna is designed for gigahertz (GHz) frequency range up to 6 GHz. The proposed frequency range accommodates the various bands between 1 and 6 GHz which lies between the wireless fidelity and Wi-Max range. The multilayer monopole patch-based microstrip antennas are gaining importance in the applications of wireless local area networks (WLAN) and wireless fidelity. The simulation of various parameters such as return loss (S11), VSWR, and radiation pattern and Bandwidth are carried out using CST software, and the screenshots of the results are given out in this chapter. The new proposed microstrip antenna has enhanced gain and bandwidth. The optimized dimensions of the geometric parameters are given in Table 11.1. There is a good agreement with the simulated results of CST software. All the comparable results of the software are achieved by simulation and approximation for the proposed design. There are different parameters of antenna which are used to examine the efficient functioning of the antenna. These are return loss (S−11), voltage standing wave ratio (VSWR), gain, bandwidth, and number of bands.

Table 11.1 Design parameters of fractal monopole antenna

S. No	Design	Dimensions $(mm)^3$				
		Substrate	Ground	Patch		
				Patch-1	Patch-2	Upper patch
1.	Design-I monopole patch	40 × 45 × 1.6	40 × 45 × 0.0635	21 × 23 × 0.0635	30 × 28 × 0.035	20 × 16 × 0.035
2.	Design-II monopole Minkowski fractal	40 × 45 × 0.0126	40 × 13 × 0.0635	21 × 23 × 0.0635	30 × 28 × 0.035	20 × 16 × 0.035

Design—III Monopole Modified Minkowski Fractal	*Antenna dimension*	*Dimension length (L) × Width (W) × height(h)mm^3*
	Substrate (s) RT/duroid	40 × 45 × 0.125
	Ground (Gl) copper	4D × 12 × 0.0635
	P1[Rectangular pole copper]	21 × 23 × 0.0635
	P2[Rectangular pole SU8]	30 × 25 × 0.12
	P1 *feed line* (f_L) P2 *feed line* (f_L)	**(0.6:0.2) × 10 × 0.0635** 1 × 10 × 0.12
Design—IV Monopole Modified Minkowski Fractal	*Antenna dimension*	*Dimension length (L) × Width (W) × height(h)mm^3*
	Substrate (s) RT/duroid	40 × 45 × 0.125
	Ground (Gl) copper	40 × 2 × 0.0635
	P1 [Rectangular pole copper]	21 × 23 × 0.0635
	P2 [Rectangular pole SU8]	30 × 28 × 0.12
	FL [SU8]	1 × 28 × 0.12

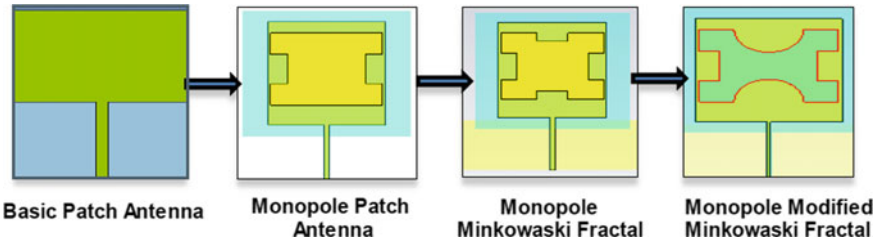


Fig. 11.2 Design of evaluation of different antenna

(a) Design-I Monopole Patch

Figure 11.3 shows the complete result with all three bands. The three different return loss (S-11) are obtained in between range 1 and 6 GHz. The S-parameter shows the RF energy propagated in the multi-port network. The *X*-axis shows the frequency range, and the Y-axis shows the return loss in dB. The result shows that there are two different bands between 1 and 6 GHz. First discuss a band that is obtained in the range of 2.475–3.017 GHz; in this band the obtained resonating frequency is 2.68 GHz; the obtained return loss is −14.244 dB; second one is at 5.2 GHz; the obtained return loss is −17.95. Totally, there are two narrow bands obtained in this design, bandwidth 22% and 23.52%. Since the bandwidth of both bands are under 50% that is why both are narrowband and proposed design 1 lies in the category of narrowband antenna (Figs. 11.4, 11.5, 11.6 and Table 11.2).

The gain of the proposed design at 5.5 GHz is near to 3.65 dB.

The performance of the above design is not up to the mark, in terms of return loss (S−11) and VSWR. The design requires enhancement and changes in the geometry of the fractal antenna (Fig. 11.7).

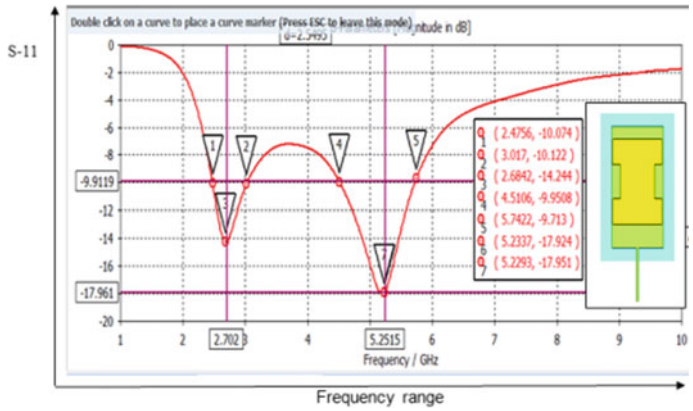


Fig. 11.3 Return loss (S−11) of proposed antenna combined two bands

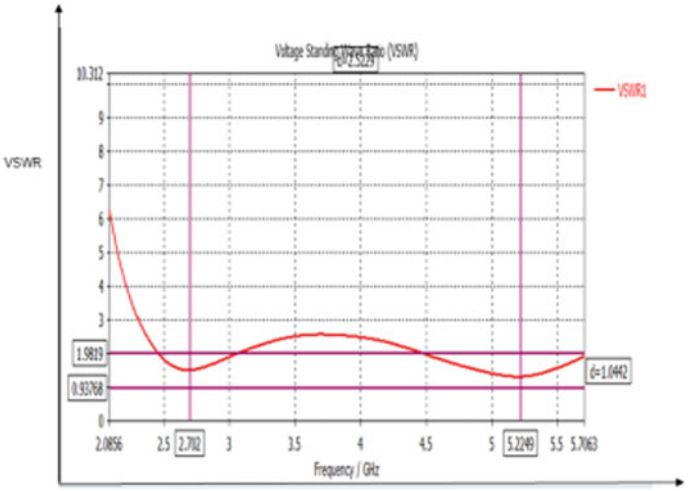


Fig. 11.4 VSWR

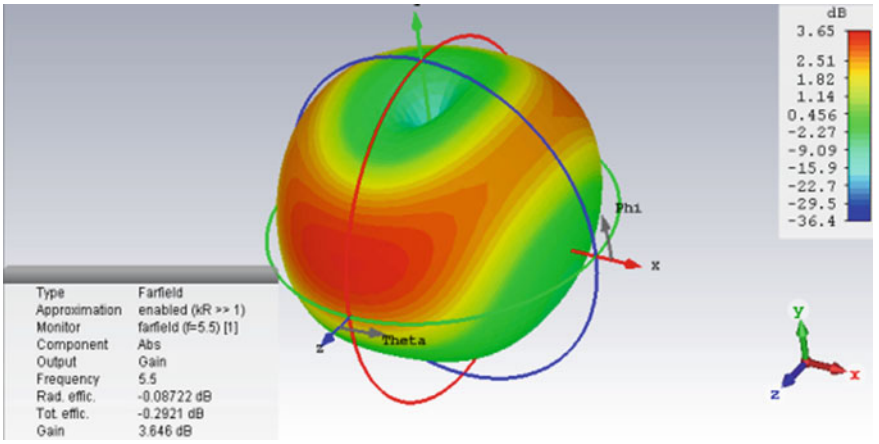


Fig. 11.5 3D diagram of the proposed design ($f = 5.5$ GHz)

(b) Design-II Monopole Minkowski fractal

A single wideband antenna is obtained in the range of 2.475−4.99 GHz; in this wideband, two resonate frequencies are obtained, first at 2.75 GHz where the return loss is −20.23 dB and the second one at 4.5 GHz where the return loss is −15.404. The overall percentage bandwidth of the proposed design 2 is 69.61%. The bandwidth of both bands is above 50% that is why design 2 exist in the wideband region.The output of the proposed design is shown in Fig. 11.8 wherein the VSWR at 2.7 GHz

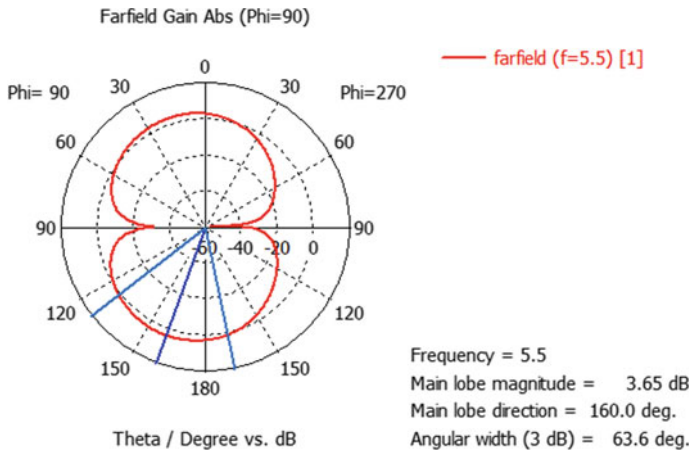


Fig. 11.6 Far-field pattern of the ($f = 5.5$ GHz)

is 1.2, and at 4.5 GHz, the VSWR is 1.4. At both resonating frequencies, the VSWR always lies under 2 (Figs. 11.9, 11.10, 11.11 and 11.12 and Table 11.3).

(c) Design—III Monopole Modified Minkowski Fractal

In design III, a parametric study is carried out of the different antenna parameters. In this design, we have analyzed the microstrip feed line in the range from 0.2 to 0.5 mm. In this range, the optimum result is obtained at a 0.2 mm feed line (f_l). Table 11.4 shows the different outcomes at different frequencies.

From the above results, it is observed that when the width of the feed line deceases, the return loss incenses and a similar effect on bandwidth also reduces with reducing the bandwidth feed line.

The output of proposed antenna design III for VSWR is good in overall bands. The VSWR at different frequencies from 2.3 to 5.9 GHz always lies under 2.5 (Figs. 11.13 and 11.14).

The gain of the proposed design at 5.5 GHz is near to 4.07 dB.

In the above Table 11.5, the comparison of all three designs having different advantages is carried out; in the case of design I, it contains two separate bands, design II contains a single band with 69.61% bandwidth and two resonant frequencies. But the design III is better as compared to all three designs (Table 11.6).

11.5 Conclusion

This paper represents the design and analysis of microstrip patch antenna by modifying Minkowski geometry for a frequency band of 1 to 6 GHz. The simulation is done by using CST studio suite 2016. The return loss is −37.11 dB for $fL = 0.3$ and

Table 11.2 Analytical results for antenna—design-I

Method	Shape	Feed technique	Range	S–parameter (dB)	VSWR	Bandwidth (%)	No. of bands
Proposed Method—1	Multilayer proposed—1	Microstrip feed	1–6	2.68 GHz= −14.35dB 5.2 GHz= −17.98dB	1.45 1.26	22 23.52	Two Narrow Bands

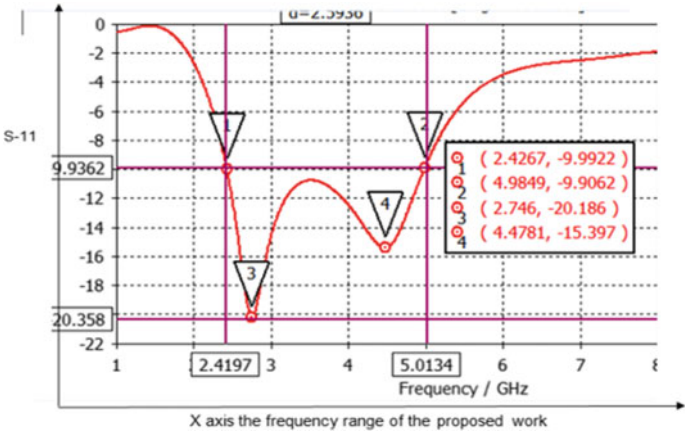


Fig. 11.7 Return loss (S−11) of proposed design II (Single wideband)

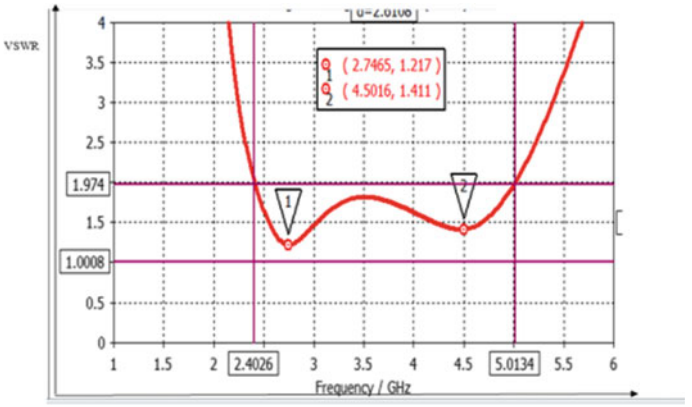


Fig. 11.8 VSWR

−59 dB for $fL = 0.2$. The VSWR is 1.02 ($fL = 0.3$), and bandwidth is 77.1% ($fL = 0.3$). This presented work shows the design of a multilayer monopole patch antenna with a Minkowski fractal. The proposed design shows good results in terms of basic antenna parameters such as VSWR, gain, return loss, and bandwidth. The proposed antenna shows a wideband and covers Wi-Fi and Wi-Max ranges whose frequencies are between 1 and 6 GHz. The overall gain of the proposed antenna is above 4db. Also, it shows the good result in terms of return loss that is (S−11) −37.11 dB as well as VSWR that is 1.02, and important parameter is percentage bandwidth which is 77.01%.

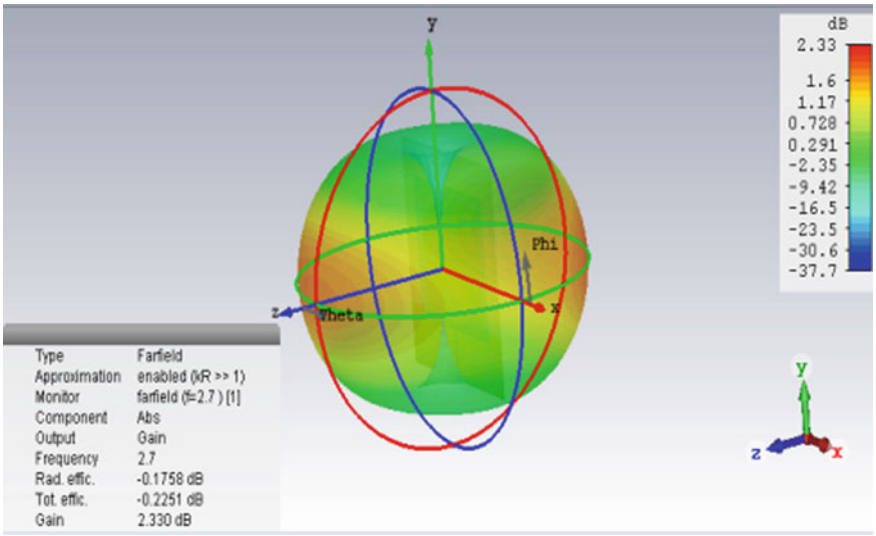


Fig. 11.9 3D diagram of the proposed design ($f = 2.7$ GHz)

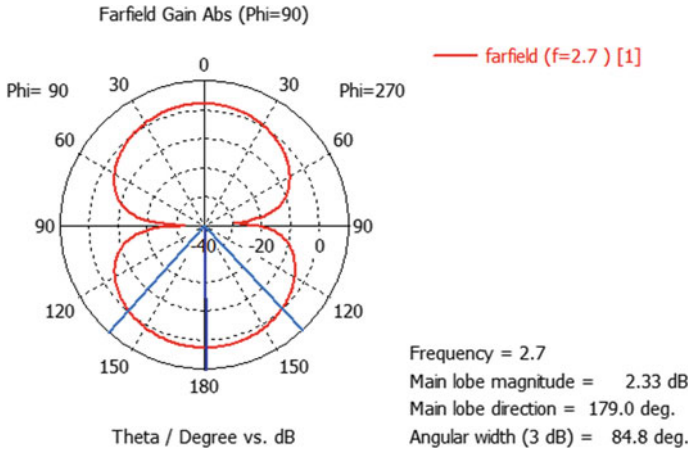


Fig. 11.10 Far-field pattern

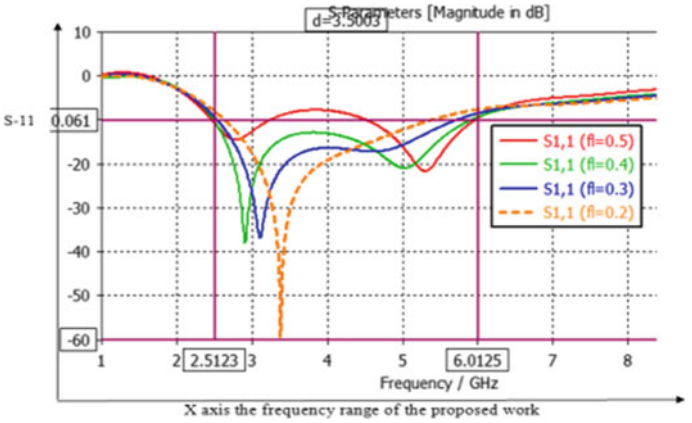


Fig. 11.11 Return loss (S-11) antenna design III (Single wideband)

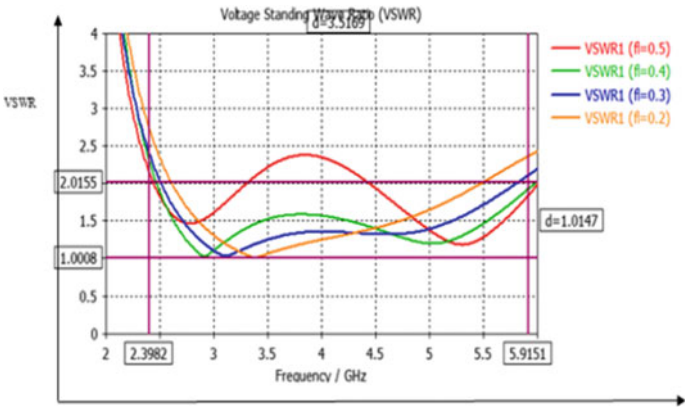


Fig. 11.12 VSWR

Table 11.3 Analytical results for antenna—design-II

Method	Shape	Feed technique	Range	S-Parameter (dB)	VSWR	Bandwidth	No. of band
Proposed Method-2	Multilayer proposed -2	Microstrip feed	1–6	2.7 GHz = −20.37dB 4.5 GHz = −15.58dB	1.2 1.4	69.61%	Single Wide Band

Table 11.4 Analytical results for antenna—design-III

Parameters feed line (f_l).	Resultant return loss (S-11) dB	Bandwidth (B.W.) %	VSWR
Fl = 0.4	2.9 GHz = −38.56	82.35% (2.48–5.91)	1.01
Fl = 0.3	3.1 GHz = −37.11	77.10% (2.55−5.75)	1.02
Fl = 02	3.7 GHz = −59.36	67.8% (2.63−5.4)	1.01

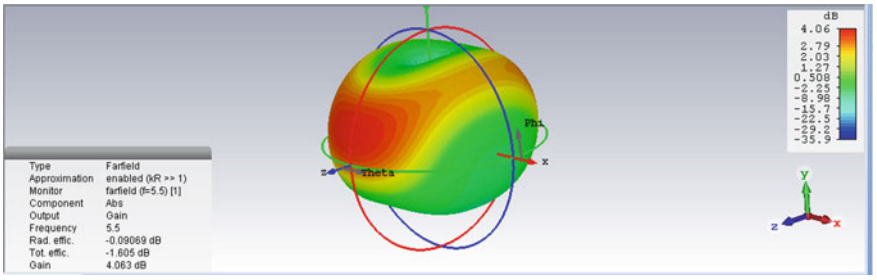


Fig. 11.13 3D pattern ($f = 5.5$ GHz)

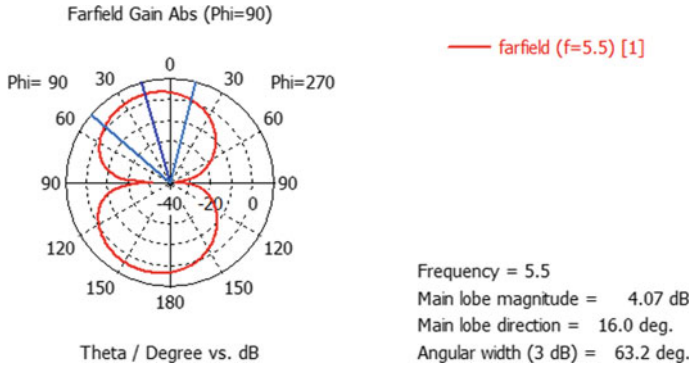


Fig. 11.14 2D pattern ($f = 5.5$ GHz)

Table 11.5 Comparative analysis between different design of antenna

Shape	Feed technique	Range	S—parameter (dB)	VSWR	Band width (%)	No. of band
Multilayer proposed—1	Microstrip feed	1–6	2.68 GHz = −14.35dB 5.2 GHz=−17.98dB	1.45 1.26	22 23.52	2 narrow
Multilayer proposed—2	Microstrip feed	1–6	2.7 GHz =−20.37dB 4.5 GHz = −15.58dB	1.2 1.4	**69.61**	Single wideband
Multilayer proposed design—3	Microstrip feed line	1–6	**3.3 GHz =−59.57dB**	1.1	67.8	Single wideband
Multilayer proposed design—3	Microstrip feed line	1–6	**3.1 GHz =−37.11**	**1.02**	**77.1**	Single wideband

Table 11.6 Comparison on the basis of S−11 and number of bands

S. No	Year	Antenna (Shape/ modification)	Feed technique	Range	S-Parameter	No. of bands
(a)	2018	Proposed multilayer patch antenna	Microstrip feed	1–6 GHz	**3.3 GHz = −59.57dB**	1 (Wideband)
(b)	2017	Flexible bow Tie antenna	CPW feed	1–6 GHz	−35 dB (Wideband bandwidth 1.79 GHz)	1

Appendix 'A'

Antena Design and Parameters

For design purposes, the Rogers RO4003C substrate is used which has a dielectric constant of 3.38. Frequency −1 to 6 GHz.

$$W = \frac{C}{2fo\sqrt{\frac{\varepsilon r+1}{2}}} \quad \varepsilon_{\text{reff}} = \frac{\varepsilon r+1}{2} + \frac{\varepsilon r-1}{2}\left[1+12\frac{h}{w}\right]^{-1/2}$$

$$\text{L}_{\text{eff}} = \frac{c}{2\ fo\sqrt{\varepsilon_{\text{reff}}}} \quad \frac{\Delta L}{h} = \frac{0.412(\varepsilon_{\text{reff}}+0.3)\left(\frac{w}{h}+0.264\right)}{(\varepsilon_{\text{reff}}-0.258)\left(\frac{w}{h}+0.8\right)}$$

$$L = L_{\text{eff}} - 2\Delta L\ Wg = 6h + L\ Z_o = 50\ \Omega\ \text{h} = 1.6\ \text{mm}.\ \lambda_o = c\,/\,f$$

$$\frac{w}{h} = \left(\frac{\exp(H')}{8} - \frac{1}{4\exp(H')}\right)^{-1}$$
$$H' = \frac{Z_0\sqrt{2(\varepsilon_r+1)}}{119.9} + \frac{1}{2}\left(\frac{\varepsilon_r-1}{\varepsilon_r+1}\right)\left(l_n\frac{\pi}{2} + \frac{1}{\varepsilon_r}l_n\frac{4}{\pi}\right)$$
$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r+1}{2}\left[1 - \frac{1}{2H'}\left(\frac{\varepsilon_r-1}{\varepsilon_r+1}\right)\left(l_n\frac{\pi}{2} + \frac{1}{\varepsilon_r}l_n\frac{4}{\pi}\right)\right]^{-2}$$

References

1. Liu, P., Zou, Y., Xie, B., Liu, X., Sun, B.: Compact CPW-fed tri-band printed antenna with meandering split-ring slot for WLAN/WiMAX applications. IEEE Antennas Wireless Propag. Lett. **11**, 1242–1244 (2012)
2. Ramarao, B.: CPW-fed monopole antenna with l shaped and stair shape slot for dual-band WLAN/WiMAX applications. In: International Conference on Innovations in Engineering and Technology, December 2013, vol. 13. pp. 65–75. (2013)
3. Moosazadeh, M., Kharkovsky, S.: Compact and small planar monopole antenna with symmetrical L-and U-shaped slots for WLAN/WiMAX applications. IEEE Antennas Wireless Propag. Lett. **13**, 1536–1225 (2014)
4. Shao, S., Kiourti, A., Burkholder, R.J., Volakis, J.L.: Broadband textile-based passive UHF RFID tag antenna for elastic material. IEEE Antennas Wirel. Propag. Lett. **14**, 1385–1388 (2015)
5. Hamouda, Z., Wojkiewicz, J.-L., Pud, A.A., Kone, L., Belaabed, B., Bergheul, S., Lasri, T.: Dual-band elliptical planar conductive polymer antenna printed on a flexible substrate. IEEE Trans. Antennas Propag. **63**(12), 5864–5867 (2015)
6. Quarfoth, R., Zhou, Y., Sievenpiper, D.: Flexible patch antennas using patterned metal sheets on silicone. IEEE Antennas Wirel. Propag. Lett. **14**, 1354–1357 (2015)
7. McKerricher, G., Titterington, D., Shamim, A.: A fully inkjet-printed 3-D honeycomb-inspired patch antenna. IEEE Antennas Wirel. Propag. Lett. **15**, 544–547 (2016)
8. Higashi, H., Rutkowski, T.M., Tanaka, T., Tanaka, Y.: Multilinear discriminant analysis with subspace constraints for single-trial classification of event-related potentials. IEEE J. Selected Topics Signal Process. **10**(7), 1295–1305 (2016)
9. Ali, T., Khaleeq, M.M., Biradar, R.C.: A multiband reconfigurable slot antenna for wireless applications. Elsevier Int. J. Electron. Commun. **84**, 273–280 (2018)
10. Singh, H.K., Paik, H.: A compact single layer differentially-fed monopole antenna for UWB applications. In: International Conference on Communication and Signal Processing, July 28–30, pp 1–4. (2020)

Chapter 12
Fake News Detection Using Text Analytics



Uma Maheshwar Amanchi, Nithesh Badam, and Rama Lakshmi Elaganti

Abstract Fake news is a form of news consisting of false statements from the real ones spread via news media or online social media. In this paper, we aim for the fake news detection model which is capable of detecting the fake news from large amounts of data that are daily produced on online platforms. The approach for our model is a machine learning technique which is text analysis and for classifying fake news we have used k means clustering. Using the data preprocessing, classification, and topic modeling we get topics from the article, and they are compared with legitimate news. We modeled a framework named Fake News Detection (FND) which is used to classify the news articles. By streaming detection of fake information, we can control false or inaccurate content.

12.1 Introduction

Fake news is considered as the dangerous thing happening in today's world and previous fake news detection techniques are inadequate in classifying the fake news from the vast amount of data that daily publishing on online platforms. So we have decided to create a fake news detection system that deliberately classifies the fake news from a large amount of data with better credibility evaluation and accuracy. In our paper, we decided the text analytics approach to detecting fake news. This approach is used in various applications and gives better results in classification projects. The text analytics approach is the machine learning model which is used to analysis the meaning of the text. (Manzoor et al. 2019) In the first phase, the data is formed into topic clusters using classification module and formed clusters based on topics using topic modelling. Next the topic clusters from the testing are formed into a list and compared with the real news topics. According to the credibility score of

U. M. Amanchi (✉) · N. Badam · R. L. Elaganti
Chaitanya Bharathi Institute of Technology, Hyderabad, India

R. L. Elaganti
e-mail: eramalakshmi_it@cbit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_12

an article we classify the article as real or fake (Hiramath & Deshpande 2019). This approach is able to give better credibility and classification accuracy than previous researches we have searched.

12.2 Proposed System

In our proposed model the model training framework has several stages of the following: We proposed a topic-based approach to detecting fake news. In first module, from the legitimate news we extract topics, then the topics are classified into various clusters. Every cluster is formed based on similar news which are common in topic. Each topic clusters contain large number of topic which are legitimate news.

12.2.1 Module 1

- At first, we collect the legitimate news and we extract topics from collected training dataset, next the news is categorized into selected number of clusters based on topics.
- Each cluster which is formed using clustering are centered on common topics of a news article.

We collected the training dataset from Kaggle website for model training framework, which are regarded as legitimate news because it is a trustworthy website by large audience. In detecting the fake news we created a framework fake news detection system (FND) undergo three modules and data from testing datasets are formed as topic clusters. Then the topic clusters are which are from the testing data is compared against the topic article from training data and using the threshold we calculate the credibility of testing articles.

12.2.2 Module 2

- In this module, the article which is to be checked is converted into clusters according to topics, the we validate this clusters with the topic clusters formed from legitimate news.

To calculate the credibility of news article we compare the topic and clusters and based on the threshold values we will estimate whether the news is real or deceptive news.

12.3 Implementation

In this section, first, we describe the model training framework which tells the detailed information about how the legitimate news datasets undergo various stages to form a well-defined topic-based cluster which is legitimate to classify from the fake ones. Then the stages of modern training framework describe data preprocessing where the data is cleaned. Second feature extraction to obtain the features from data to form clusters. Finally, the topic modeling approach is used to form clusters that are surrounded by common news topics. The required evaluation metrics calculations are described in detail with formulas.

By The design and development of the model training framework are as shown in Fig. 12.1. The model training framework is driven by a ground-truth knowledge base comprised of legitimate-news topic clusters. The model-training framework drives the functioning of how it will create clusters based on topics such that news articles in the same cluster share a set of similar topics. Topics classified in separate clusters are different from each other cluster.

This framework guides the design and development of detecting fake news. The model training framework describes the detailed procedure of how legitimate news is classified according to similar news topics. This will act as a ground truth model for converting any news which is to be verified. The legitimate news is free from stemming and stop words when it undergoes through the framework. All the datasets are first undergone through model training framework and comparison is done to detecting the news which is fake.

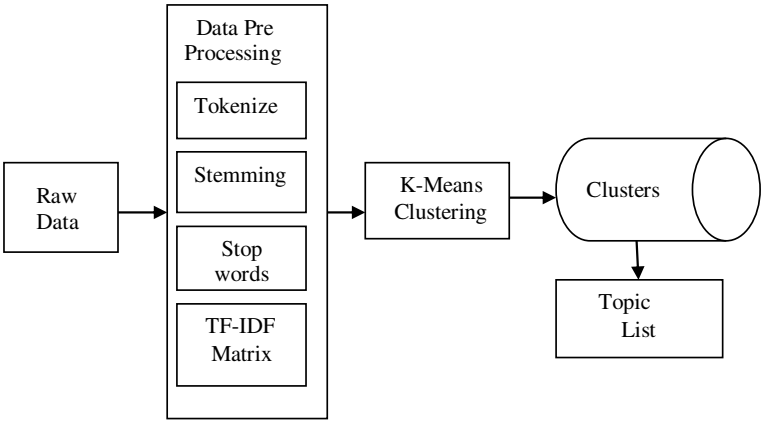


Fig. 12.1 Model training framework

12.3.1 FND (Fake News Detection System)

They are various stages involved in the FND system of detecting fake news from various sources. The input from any source undergoes many operations as described above in the model training framework which are data preprocessing, feature extraction, clustering, and topic modeling. The FND system checks the news which should be verified as fake or real news from the dataset. It compares the news which should be verified and the news will undergo the model training framework and compared with the legitimate news clusters from the model training framework.

In the fake news detection system, the input from any source is converted into the CSV file format. The data will undergoes data preprocessing to free from all data errors. Then the data will undergo TF-IDF vectorization to behave as a feature to form as a cluster with similar topics. Similar topics are under a single cluster. Each cluster consists of various large numbers of topics. The comparison of a large number of topics is not easy to handle by the clustering technique. The topics are formed as a cluster based on topics using topic modeling Similar topics are surrounded by common shared topics using topic modeling (Ibrishimova 2019). We have used the LDA model to separate the topics and the clusters are compared against the legitimate news topic cluster to verify whether the news is fake or real.

Figure 12.2 describes FND system consists of three modules each module i.e., data preprocessing, classification module, verification module. We collected the two datasets which are training and testing from Kaggle for model training, which are real and fake news articles respectively. In our training dataset there are 2000 articles and 100 articles in testing dataset. They also provide fake news dataset where the news should be verified. The data from the training dataset will undergo various techniques to be free and using the clustering algorithm the similar news formed as

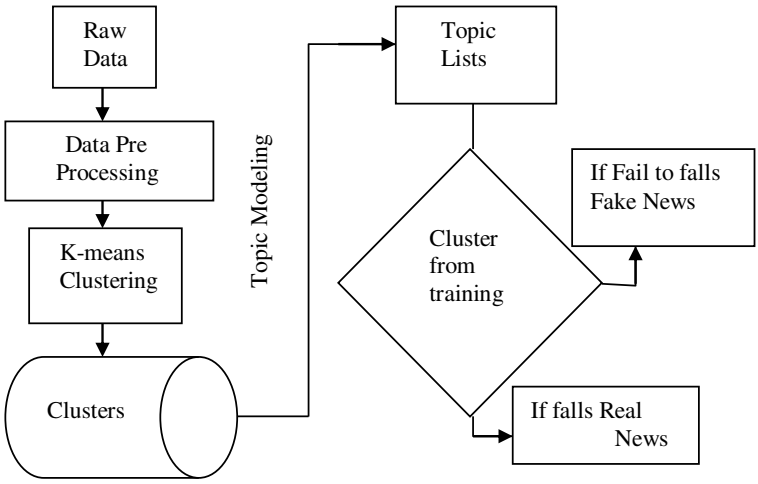


Fig. 12.2 Fake news detection system

a single cluster. Various clusters are formed from the clustering algorithm and the clusters are converted to topic clusters where the article with similar topic information into single topic cluster. Topics from the testing dataset compared against the topics of training dataset based on the threshold we calculate the credibility of the news article.

12.3.2 Data Preprocessing

Data Preprocessing is the process of analyzing the data that has not been screened for such problems that can produce misleading results. The raw data from the various sources are saved as ground-truth and fake news database. Data undergoes various stages of data cleaning which can give the data free from errors. The various techniques involved in data preprocessing are the following:

Tokenization: It is a very common task in NLP while dealing with the characters, the process of chopping a character into pieces (which are known as tokens) called tokenization.

Stemming: The process of reducing the word into its root word is called stemming.

Lemmatization: It is a process of grouping the inflected form of words so they can be analyzed as a single term.

Stop Words: Those are the word which is filtered on and or before data preprocessing techniques applied on data.

12.3.3 TF-IDF Vectorization

TF: Term Frequency
IDF: Inverse Document Frequency
TF-IDF = TF*IDF
TF = (Number of times term the particular word appears an article)/(Number of words in an article)
IDF = log(S/s)
S is the total number of articles in the datasets, and s is the number of articles the particular word has appeared in the dataset. Clustering can be formed by using only numerical values, hence TF-IDF values for each word is calculated and clustering is formed by using the values.

12.3.4 Clustering

In classification module, we have used an unsupervised clustering algorithm to classify similar news from large datasets. We have used the k means clustering algorithm

to classify the fake news from the real news. In our model, we have formed 50 clusters each cluster has centers and surrounded by similar news. The clusters formed from the algorithm may contain a large number of news and handling a large number of news in a single cluster difficult and the articles undergo topic modeling technique to form topic-based clusters.

In the implementation of k means clustering first (Mahid 2018), we assume that we get an output of k clusters. Then we randomly select the k number of items and using euclidean distance each k item is surrounded with similar news with a centroid. Then the process of euclidean distance is repeated until the appropriate news is fallen into the particular cluster.

12.3.5 Topic Modeling

In this paper, the topic modeling is used to create the topic based clusters. We have used LDA topic modeling technique for the topic-based representation of clusters. The clusters are formed based on topics and each topic contain similar information. The topics from the training and testing are stored in the list and the topics of article which is to be checked is compared against the topics of training and calculate the credibility of article.

12.3.6 Evaluation Credibility

The information of articles from testing data are used to check the credibility of system. The credibility of test news article is calculated through a function which described whether the news is fake topic and real topic, using the threshold we display the false credibility of an article.

$$g(\alpha) = \frac{\sum_{j=1}^{w_i} (f_E(\mathrm{e}_i^j))}{w_i} \tag{12.1}$$

Thus, we have where α is the news article which is to be tested, w_i is the number of topics generated from the test news, f_E (e^{i^j}) is the Boolean valued function.

12.3.7 Performance Evaluation

- $\mathrm{d}_{r\to r}$ be the number of news which are true and truly detected as true information.
- $\mathrm{d}_{r\to f}$ is the number of news which are true and falsely detected as false information.

- $d_{f \to r}$ is the number of news which are false and falsely detected as true information.
- $d_{f \to f}$ is the number of news which are false and truly detected as false information.

Accuracy Rate (*A*)

$$A = \frac{d_{r \to r} + d_{f \to f}}{d_{r \to r} + d_{r \to f} + d_{f \to r} + d_{f \to f}} \tag{12.2}$$

Precision (*P*)

$$P = \frac{d_{f \to f}}{d_{r \to f} + d_{f \to f}} \tag{12.3}$$

Recall (*R*)

$$R = \frac{d_{f \to f}}{d_{f \to r} + d_{f \to f}} \tag{12.4}$$

F-Score (*F*)

$$F = \frac{2 \times P \times R}{P + R} \tag{12.5}$$

12.4 Results

12.4.1 Cluster Evaluation

Topics from the trained datasets are compared against the training dataset based on threshold valued the testing article is classified as legitimate or deceptive news respectively is shown in Fig. 12.3.

12.4.2 Accuracy of FND System

Figure 12.4 Illustrates the proposed approach for detecting the fake news from large amount of data which is generated from social media using the machine learning techniques. Our model got better classification accuracy than various other approaches to detecting fake news and accuracy about 88%.

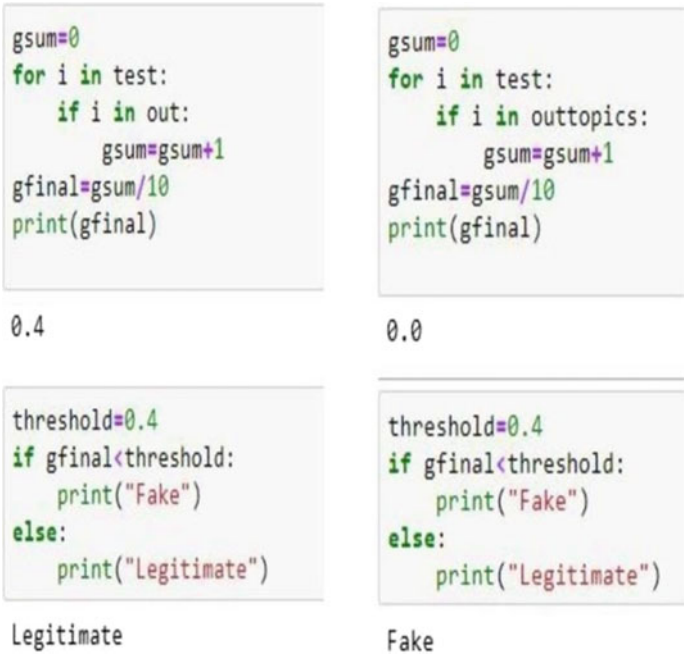


Fig. 12.3 Evaluation of fake news

Fig. 12.4 Accuracy of FND systems

```
c=data['Stance']
ct=0
for i in range(0,904):
  if l[i]==1 and c[i]=='Fake':
    ct+=1
  if l[i]==0 and c[i]=='Legitimate':
    ct+=1
print('Accuracy:',ct*100/904)
```

```
Accuracy: 87.83185840707965
```

12.5 Conclusion

Detecting fake news has become a crucial role in everyday life, various techniques are lack in detecting the deceptive news from large data produced every day from various social platforms. In this paper, we have proposed a model that is used to detect fake news continuously produced on online media. We have created a text

analytics-based approach system mainly focused on representing the news as topic-based representation and the system is used to classify the news as deceptive or real. We have used three modules in classifying the news which is to be verified those three modules make various changes to the data and made a topic-based representation of news. Topics from the testing data is compared against the topics from training data which are also derived using three modules used to classify the news. Each topic is surrounded by similar news and based on threshold value the news is classified as fake news. Our model got better classification accuracy than various other approaches to detecting fake news.

References

1. Hiramath, C.K., Deshpande, G.C.: Fake news detection using deep learning techniques. In: Science Direct on Multimedia, Chikmagalur, India, pp. 411–415 (2019)
2. Ibrishimova, M.D., Li, K.F.: A machine learning approach to fake news detection using knowledge verification and natural language processing. In: Barolli, L., Nishino, H., Miwa, H. (eds.) Advances in Intelligent Networking and Collaborative Systems. INCoS 2019. Advances in Intelligent Systems and Computing, vol 1035. Springer, Cham (2020)
3. Manzoor, S.I., Singla, J., Nikita: Fake news detection using machine learning approaches: a systematic review. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, pp. 230–234 (2019)
4. Mahid, Z.I., Manickam, S., Karuppayah, S.: Fake news on social media: brief review on detection techniques. In: Science Direct, Subang Jaya, Malaysia, pp. 1–5 (2018)

Chapter 13
An Extreme Learning Machine-Based Model for Cryptocurrencies Prediction



Sarat Chandra Nayak, B. Satyanarayana, Bimal Prasad Kar, and J. Karthik

Abstract Cryptocurrency, in a short period of time has got wide popularity and considered as an investment asset. Prediction of cryptocurrency is a recent area of research interest and budding fast. The price trend of cryptocurrency behaves arbitrarily and fluctuates like other stock markets due to inherent volatility. Though few computational intelligence methods are available, sophisticated methodologies for accurate prediction of cryptocurrency are still lacking and need to be explored. Extreme learning machine (ELM) is a faster and better learning method for neural networks with solitary hidden layer and has enhanced generalization performance. This study proposes an ELM based approach for prediction of four emerging cryptocurrencies such as Litecoin, Ethereum, Ripple, and Bitcoin. The prediction ability of the proposed approach is compared with few similar methods such as RBFN, SVM, MLP, ARIMA, and LSE. From exhaustive simulation studies and comparative result analysis it is found that the ELM method performed better than others and hence can be suggested as an efficient tool for cryptocurrencies prediction.

13.1 Introduction

Cryptocurrency is an open source and decentralized digital currency based on principle of cryptography. The relevance of cryptocurrencies as an emerging market in the financial world is increasing significantly. Now days more and more financial organizations are getting concerned about cryptocurrency trading. The trading opportunities and profitability in the cryptocurrency market are studied in [1]. A survey on systems of cryptocurrency is elaborated in [2]. The characteristics of this

S. C. Nayak (✉) · J. Karthik
CMR College of Engineering and Technology, Hyderabad 501401, India

B. Satyanarayana
CMR Institute of Technology, Hyderabad 501401, India

B. P. Kar
Gandhi Institute for Technological Advancement, Bhubaneswar 752054, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_13

market such as high availability of market data, high volatility, smaller capitalization, and decentralized control are studied in [3]. This market fluctuates like other stock markets due to inherent volatility and the confidence of investors have been reflected on it [4]. The feasibility of this market in relation to other financial markets is documented in the literature [5]. Like other stock market, cryptocurrency market prices behave arbitrarily and coupled with high nonlinearity and dynamics. Though few computational intelligence methods are available, sophisticated methodologies for accurate prediction of cryptocurrency are still lacking and need to be explored.

As the cryptocurrencies market prices are growing fast and behaves similarly as other stock market price movement, speculators as well as researchers are anxious about prediction of their market values. During last few years machine learning and statistical methodologies are used for prediction of Bitcoin [6, 7]. Forecasting accuracy of a neural model is greatly subjective to the network magnitude and learning method. Gradient descent learning is a common method for neural network training. However, suffering from sluggish convergence rate, imprecise learning and prone to local minima are the general drawbacks of this technique. It adds computational overhead to the model [8]. To overcome these, ELM was proposed in [9]. It chooses the input-hidden weights at random. The hidden-output weights are determined analytically in contrast to iterative fine-tuning. Numerous researches and experimentations are done for diverse applications using ELM [10]. It is a faster and better learning method for networks with solitary hidden layer and has enhanced generalization performance.

With objective of achieving improved prediction accuracy, here we use ELM for training of a neural network having single hidden layer. The trained network is then used to extrapolate the historical prices of the four cryptocurrency and predict the next day price. Contrast to conventional method of segregating the dataset into two parts (i.e. training and testing), we followed a moving window method to train and test pattern generation. To access the performance of ELM based approach, five other models such as RBFN, SVM, MLP, ARIMA, and LSE are developed in similar fashion. Prediction accuracies of all the methods are accessed in terms of MAPE, ARV, and computation times from four cryptocurrency datasets.

The remaining parts such as methods and materials are described in Sect. 13.2, the ELM-based forecasting method is explained in Sect. 13.3, and simulation results are summarized in Sect. 13.4 followed by concluding remarks.

13.2 Methods and Materials

The methods such as ELM, and proposed ELM-based forecasting and materials such as cryptocurrency datasets collected from online sources are presented in this section.

13.2.1 ELM

As stated earlier, to overcome the limitations of gradient descent learning ELM was proposed [9]. It was proposed for training networks with one hidden layer. It considers random weight and bias for hidden nodes. As an alternative to iterative tuning the output weights are determined through generalized inverse function on the outputs from the hidden layer. For an input vector x_j, weight vector between input neuron and ith hidden neuron $w_i = [w_{i1}, w_{i2}, \cdots, w_{iN}]^T$, $(i = 1, 2, \cdots, N_h)$ and output weights linking the ith hidden neuron and the output neurons $\beta_i = [\beta_{i1}, \beta_{i2}, \cdots, \beta_{iN}]^T$, the output vector O_j is computed as:

$$O_j = \sum_{i=1}^{N_h} \beta_i * f\left(w_i x_j + \text{bias}_i\right), \; j = 1, 2, \ldots, N \tag{13.1}$$

Here N_h is the hidden layer size. The output weight vector β_i obtained by solving $H\beta = Y$, where:

$$H(w_i, x_i, b_i) = \begin{bmatrix} f(w_1 x_1 + b_1) & \cdots & f\left(w_{N_h} x_1 + b_{N_h}\right) \\ \vdots & \ddots & \vdots \\ f(w_1 x_N + b_1) & \cdots & f\left(w_{N_h} x_N + b_{N_h}\right) \end{bmatrix}_{N \times N_h} \tag{13.2}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{N_h}^T \end{bmatrix}_{N_h \times m} \qquad Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times m}$$

In real cases, the inequalities $N_h \ll N$ holds true. So, H is a non-square matrix and may be non-singular in the majority of cases. Therefore, a combination (w_i, b_i, β_i) satisfying Eq. (13.2) may not exist. So, the network can be trained by finding the least square minimum norm solution $\hat{\beta}$ of (13.2) as follows:

$$\left\| H\hat{\beta} - Y \right\| = \min_{\beta} \|H\beta - Y\| \tag{13.3}$$

The minimum norm least square solution of Eq. (13.2) is calculated as in Eq. (13.4). Here H^+ is the pseudo inverse or Moore–Penrose inverse of H.

$$\hat{\beta} = H^+ Y \tag{13.4}$$

13.2.2 Proposed ELM-Based Prediction

A pictorial view of ELM-based forecasting of cryptocurrency is depicted in Fig. 13.1. Here, the base model is a neural network having single hidden layer. The network is trained by ELM. The input-hidden weights are assigned with random values. The output weights are computed as per Eq. 13.4. We created the input patterns for the model using moving window from the original series. For example one train/test set is formed as follows.

$$\underbrace{\begin{matrix} x(i) & x(i+1) & x(i+2) \\ x(i+1) & x(i+2) & x(i+3) \\ x(i+2) & x(i+3) & x(i+4) \end{matrix}}_{\text{Training data}} \underbrace{\begin{matrix} \vdots x(i+3) \\ \vdots x(i+4) \\ \vdots x(i+5) \end{matrix}}_{\text{Target}}$$

$$\underbrace{x(i+3)\ x(i+4)\ x(i+5)}_{\text{Test data}} \underbrace{\vdots x(i+6)}_{\text{Target}}$$

The patterns are then normalized to scale the data into the same range for each input feature in order to minimize bias [11]. We used tanh estimator method as in Eq. 13.5 which is an efficient scheme to standardize time series data. The mean and standard deviation of the moving window are represented as μ and σ respectively.

$$\hat{x} = 0.5 * \left(\tanh\left(\frac{0.01 * (x - \mu)}{\sigma} \right) + 1 \right) \tag{13.5}$$

After normalization the patterns are inputted sequentially to the network. The hidden neurons calculate the product of input and associated weight adds a bias and

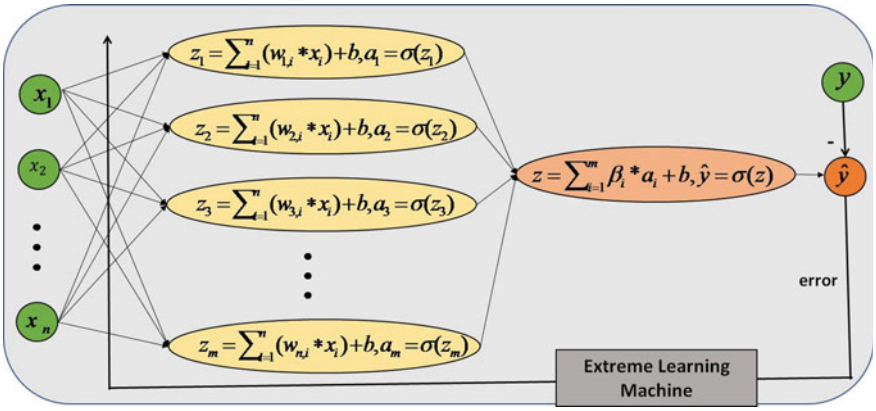


Fig. 13.1 ELM-based forecasting

Table 13.1 Statistical summary of four cryptocurrencies

Statistic	Bitcoin	Litecoin	Ethereum	Ripple
Minimum	68.4300	1.1600	0.4348	0.0028
Mean	$1.4857e^{+03}$	20.4966	147.7843	0.0984
Median	482.8100	3.9100	12.0200	0.0079
Variance	$8.7573e^{+06}$	$2.2240e^{+03}$	$6.9765e^{+04}$	0.1028
Maximum	$1.9497e^{+04}$	358.3400	$1.3964e^{+03}$	3.3800
Std.dev	$2.9593e^{+03}$	47.1594	264.1308	0.3206
Skewness	3.5394	4.1627	2.38440	5.8039
Kurtosis	15.9671	21.5925	8.51220	42.9469
Corr. Co.eff	0.00030	0.00211	−0.00520	0.00027
No. of data	1760	1760	929	1662

applies sigmoid activation to compute the net output. The output unit computes the product of weights and activations from the hidden units adds a bias and applied activation. The model estimated an output ($\hat{y}$) at the output layer as in Eq. 13.6. The amount of deviation of the estimation from the target (y) is calculated as error, i.e. Equation 13.7. The minimum is the error; more closer the model estimation towards the actual, hence better is the prediction accuracy of the model.

$$\hat{y} = \sum_{i=1}^{m} \sigma(\beta_i * a_i + b) \tag{13.6}$$

$$\text{error} = \text{abs}\left(y - \hat{y}\right) \tag{13.7}$$

13.2.3 Cryptocurrency Data

The models are evaluated on experimenting cryptocurrencies historical closing prices collected from kaggle.com. The currencies are Bitcoin, Litecoin, Ethereum, and Ripple. A summary of statistics of the four datasets are given in Table 13.1. The price series are plotted in Fig. 13.3. The total number of data points on the Bitcoin, Litecoin, Ethereum, and Ripple are 1760, 1760, 929, and 1662 respectively.

13.3 Experimental Results and Analysis

To evaluate the models, we used them to forecast the prices of four cryptocurrency datasets separately. To access the capacity of the ELM method, five other methods

Table 13.2 MAPE and ARV from all models

MODEL	MAPE				ARV			
	Bitcoin	Litecoin	Ehtereum	Ripple	Bitcoin	Litecoin	Ehtereum	Ripple
LSE	0.084185	0.172655	0.078047	0.088747	0.015777	0.024671	0.019375	0.038572
ARIMA	0.060645	0.094002	0.075884	0.082855	0.008352	0.009558	0.046743	0.016345
MLP	0.047281	0.083777	0.059358	0.049708	0.007763	0.007994	0.005905	0.017437
SVM	0.039462	0.055652	0.046352	0.047638	0.006584	0.006042	0.006355	0.006854
RBFN	0.045470	0.053754	0.049535	**0.045733**	0.007563	0.007962	0.005756	0.014362
ELM	**0.032283**	**0.051376**	**0.043901**	0.047335	**0.006137**	**0.005815**	**0.003982**	**0.005333**

The best MAPE and ARV statistics values are shown in bold face.

such as SVM, MLP, RBFN, ARIMA, and LSE are developed in this study and a comparison is done. Two performance metrics such as mean absolute percentage of error (MAPE), and average relative variance (ARV) are used to measure the forecast accuracy as follows:

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|x_i - \hat{x}_i\right|}{x_i} \times 100\% \tag{13.8}$$

$$\text{ARV} = \frac{\sum_{i=1}^{N}\left(\hat{x}_{\text{i}} - x_i\right)^2}{\sum_{i=1}^{N}\left(\hat{x}_i - \overline{X}\right)^2} \tag{13.9}$$

The mean values from twenty simulations are considered as the performance of a model. The MAPE and ARV values are summarized in Table 13.2. The best statistics are highlighted in bold. It may be seen that the MAPE values from the ELM are lower than others expect Ripple dataset. For example, MAPE of ELM is 0.032283 from Bitcoin, 0.051376 from Litecoin, and 0.043901 from Ethereum. For Ripple the RBFN produced lowest, i.e. 0.047335. However, the ELM achieved lowest MAPE thrice. It achieved lowest ARV values for all four datasets. The overall performance of ELM is better than others. It also may be observed that the SVM and RBFN are placed in second and third position. Further, to show the goodness of the ELM method the predicted prices are plotted against the actual and depicted in Figs. 13.2, 13.3, 13.4, 13.5 respectively. It can be seen that the ELM predicted prices are very closer to the actual prices and following the trend accurately.

13.4 Conclusions

Prediction of cryptocurrency is a recent area of research interest and budding fast. The trend it follows behaves arbitrarily and fluctuates like other stock markets due

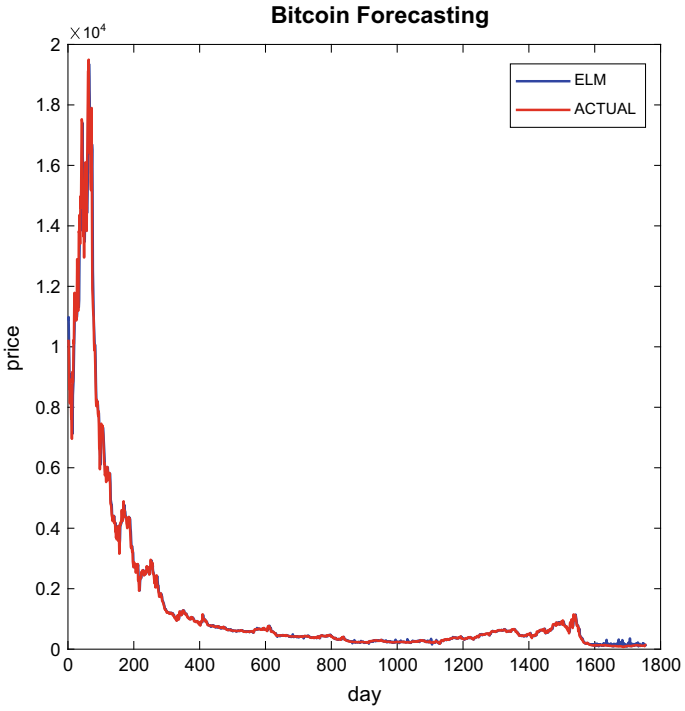


Fig. 13.2 Forecasted prices versus actual from bitcoin

to inherent volatility. Therefore, capturing the under laying dynamics of cryptocurrency prices movement is challenging. This article proposed an ELM-based forecasting approach for predicting the future price of four popular currencies such as Bitcoin, Litecoin, Ethereum, and Ripple. ELM has faster convergence rate and better generalization ability compared to gradient descent learning. Here, we used ELM for training a single hidden layer neural network. The ELM predictability is compared with that of SVM, RBFN, MLP, ARIMA, and LSE in terms of MAPE and ARV. From comparative analysis it is concluded that the ELM-based forecasting is good enough to forecast the cryptocurrencies than others. The future plan may explore other neural networks to be trained with ELM and the current approach may be applied for other financial forecasting problems.

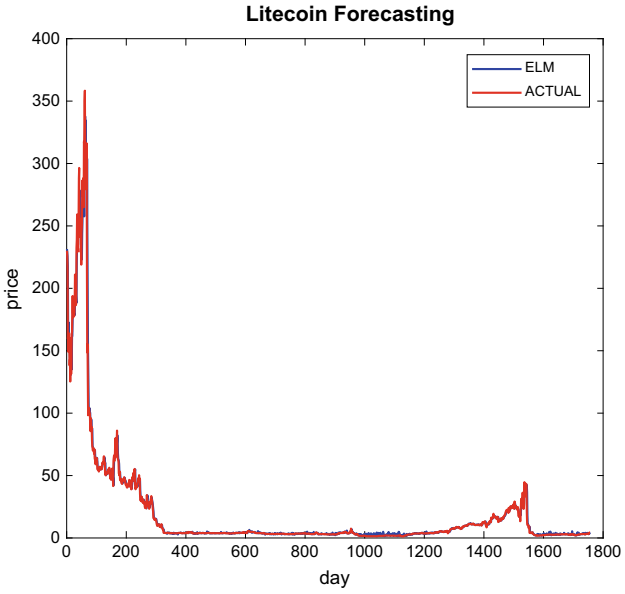


Fig. 13.3 Forecasted prices versus actual from litecoin

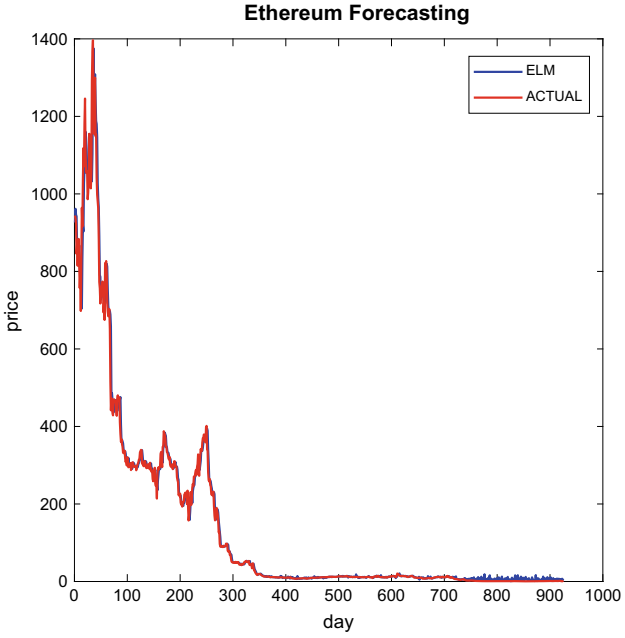


Fig. 13.4 Forecasted prices versus actual from ethereum

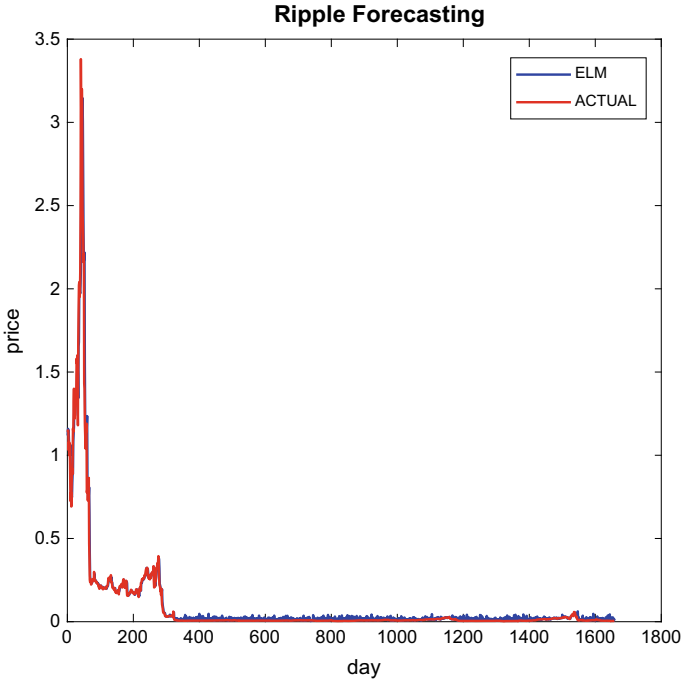


Fig. 13.5 Forecasted prices versus actual from ripple

References

1. Kyriazis, N.A.: A survey on efficiency and profitable trading opportunities in cryptocurrency markets. J. Risk Finan. Manage. **12**, 67 (2019)
2. Mukhopadhyay, U., Skjellum, A., Hambolu, O., Oakley, J., Yu, L., Brooks, R.: A brief survey of cryptocurrency systems. In: 2016 14th Annual Conference on Privacy, Security and Trust (PST), IEEE. pp. 745–752. (2016)
3. Ferreira, M., Rodrigues, S., Reis, C.I., Maximiano, M.: Blockchain: a tale of two applications. Appl. Sci. **8**, 1506 (2018)
4. Grinberg, R.: Bitcoin: an innovative alternative digital currency. Hastings Sci. Technol. Law J. **4**, 160–207 (2011)
5. Corelli, A.: Cryptocurrencies and exchange rates: a relationship and causality analysis. Risks **6**, 111 (2018)
6. Chu, J., Nadarajah, S., Chan, S.: Statistical analysis of the exchange rate of bitcoin. PloS One **10**(7) (2015)
7. Núñez, J.A., Contreras-Valdez, M.I., Franco-Ruiz, C.A.: Statistical analysis of bitcoin during explosive behavior periods. PloS One, **14**(3) (2019)
8. Fernández-Navarro, F., Hervás-Martínez, C., Ruiz, R., Riquelme, J.C.: Evolutionary generalized radial basis function neural networks for improving prediction accuracy in gene classification using feature selection. Appl. Soft Comput. **12**(6), 1787–1800 (2012)
9. Huang, G.B., Zhu, Q.Y., Siew, C.K.: Extreme learning machine: theory and applications. Neurocomputing **70**(1–3), 489–501 (2006)

10. Nayak, S.C., Misra, B.B.: Extreme learning with chemical reaction optimization for stock volatility prediction. Finan. Innov. **6**(1), 1–23 (2020)
11. Nayak, S.C., Misra, B.B., Behera, H.S.: Impact of data normalization on stock index forecasting. Int. J. Comput. Info. Syst. Indus. Manage. Appl. **6**(2014), 257–269 (2014)

Chapter 14
An Ensemble Approach for Intrusion Detection in Collaborative Attack Environment



Shraddha R. Khonde, G. Kulanthaivel, and V. Ulagamuthalvi

Abstract Data confidentiality and integrity are the main issue in modern era of Internet. Network security is compromised by intruders with help of attacks. Collaborative attack is a new attack used by intruder which is to be monitored by intrusion detection system (IDS). In this paper, we are proposing IDS which can handle collaborative attack with the help of ensemble classifiers. Proposed IDS makes use of various feature selection techniques to reduce number of feature required for attack detection. Various classifiers are used in ensemble to increase attack detection rate and reduce false alarm rate (FAR). Dataset used for experiments is UNSW NB15. Comparison of individual classifier is done with ensemble approach. Results show reduction of 5% in FAR and improvement of 2% in accuracy for ensemble approach. Ensemble approach and feature selection provide improved performance of IDS in terms of precision and recall.

14.1 Introduction

New era of Internet makes use of various technologies and architectures to provide ease and efficient data handling through network and Internet. To provide security to data transferred in Internet, network security policies should be strong enough. To maintain security of the network, various tools are used by administrators to handle outside threats. When data is compromised by any insider, it becomes difficult to identify and handle it. To handle such type of attack that is inside as well as outside attacks, an intelligent system is used known as IDS. This system monitors all activities of the network whether data is coming from outside sender or from inside sender; it is analysed and checked for security. While monitoring activities in network if any abnormal activity is observed by IDS, an alarm is generated for

S. R. Khonde (✉) · V. Ulagamuthalvi
Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India

G. Kulanthaivel
National Institute of Technical Teachers Training and Research, Chennai, Tamilnadu, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_14

administrator for taking further action on source and destination. Various types of attacks are available in this modern era which not only compromise the system, but takes whole control of system and violates the network such as *active*, *passive and collaborative*. Active attacks modifies data captured from the network and penetrate back modified data in the network. This type of attack is dangerous to the system as modified data can have different or severe impact on receiver. Passive attack does not modify data of network it only monitors network activities. Using this type of attack, intruders try to pass confidential information outside the network. Collaborative attack is a new type of attack used by intruders, nowadays, in this multimedia modern era of data transmission. This type of attack is created with the help of multiple attacks. Combinations of multiple attacks are used to generate new attacks to get entry into the system. As the signatures of most well-known single attacks are stored in dataset, it becomes cumbersome for intruder to get entry into network using single attack. This is the main reason that intruders are trying to create and use collaborative attacks to get access of network. To detect attacks, IDS makes use of various machine learning algorithms. These algorithms help IDS in classifying network traffic in normal and attack category. Attack detection by IDS is categorised as *signature-based IDS* and *anomaly-based IDS*. In signature detection, signatures stored in the dataset are used to identify malicious activities happening in the network. Signatures are nothing but footprints stored as a signature in a dataset. Each packet entering in network is matched with signature for attack detection [1]. Limitation of this detection method is datasets that are not up-to-date for modern attacks [2]. Anomaly-based IDS focussed on the behaviour of node or the system or the network to find abnormal activities. Any packet of data entering in the system is analysed according to normal behaviour. If it is deviated from normal behaviour, it is considered as an abnormal behaviour and marked as novel or anonymous attack. Little deviation in normal behaviour is considered as attack becomes limitation of this detection method which can be handled by various machine learning algorithms [3]. In this paper, we are proposing better IDS which will be able to identify or detect collaborative attack along with single attack at a time. Collaborative attack is defined as multiple attacks are happening on the network at the same time or with the difference of fraction of seconds. It can also be any new attack which is created by combining two or more existing attacks. To implement this IDS, use of various machine learning algorithms is done. To improve detection accuracy and reducing false alarm rate, ensemble of classifier is done. Experiments were done on UNSW NB15 dataset to handle modern day attacks. Collaborative attack signatures are created and updated in dataset. Feature reduction is used to reduce features for training classifiers.

14.2 Literature Survey

Network security is prime area on research nowadays because of huge usage of data sharing and Internet. This takes focus of many researchers towards various machine

learning, deep learning and data mining algorithms used to improve performance of IDS. Various datasets are available like KDD99, NSL-KDD, DARPA, UNSW-NB15, CIC-IDS2017 and many more. These all datasets are used by researchers to implement IDS for attack detection. In [4], importance and performance of supervised machine learning algorithm are explained with the help of multilayer perceptron neural network model for improving performance of IDS. Importance of ensemble method over single classifier is explained in [5]. An ensemble approach of genetic algorithm and support vector machine to improve performance of IDS is elaborated in [6] which make use of fitness function to reduce error rate. Approaches like ensemble and hybrid classifiers are used to implement IDS [7]. IDS with ensemble classifiers random forest, naïve Bayes and average one-dependence estimator is used by increasing accuracy in [8]. Another ensemble approach is presented in [9] where core vector machine is used to identify all categories of attacks of KDD99 dataset. For each category, one CVM is implemented and tested for attack detection. Results prove that this ensemble approach with CVM provides good efficiency and less computational time. Feature selection is better explored in [10, 11] as chi square, gain ratio and information gain on ensemble and single classifiers. Comparison of various machine learning algorithms according to their performance is explained in [12]. Ensemble with best classifier along with feature selection technique improves IDS performance [13]. Use of feature selection based on genetic algorithms to obtained 99.71% accuracy is elaborated in [14] over NSL-KDD dataset with 15 features. Wrapper feature selection technique can also use to receive better accuracy from classifiers [15]. Performance parameters are elaborated in [16] which help to evaluate overall performance of IDS. Hybrid architecture using both detection methods provides good performance in attack detection [17]. It helps in identifying both known and unknown attacks.

A novel framework is proposed in [18], which use SVM ensemble with feature augmentation. Dataset is transformed into new supervised dataset using marginal density function. Feature selection technique to improve performance of IDS in binary and multiclass classifiers along with hybrid approach is used to test results on NSL-KDD dataset [19]. This approach provides accuracy of 99.77% for binary and 99.63% for multiclass detection. In [20], author focusses on reducing false alarm rate in IDS to improve its performance. Performance of IDS is not only depending on the classifiers used or a feature selection technique used. Ensemble shows better performance as compared to individual classifier [21]. Survey of standard datasets is presented in [22] used for network IDS. Customization of dataset is required for maintaining performance of IDS and increasing detection rate. UNSW NB15 dataset shows better performance in terms of modern attacks [23]. This dataset consists of signatures for all modern attacks which helps IDS to detect various up-to-date attacks.

As per observations in literature survey presented above, we can say that IDS provides improved performance if classifiers are used in ensemble approach rather than individually. Feature selection is also important to reduce computation and training time of classifier. In proposed IDS, multiple supervised and unsupervised machine learning classifiers are used in ensemble to implement IDS. Experiments are conducted on UNSW NB15 dataset. To reduce number of features from dataset,

feature selection techniques are used. As none of the benchmarked dataset includes signature of collaborative attack, UNSW NB15 dataset is updated by adding signatures of collaborative attacks. For creation of signature, attack is generated using scripts and penetrated in network so that packets can be captured and signatures can be created. Performance evaluation is done by considering various parameters such as precision, recall, FAR, and accuracy of classifier. Comparison of individual classifier is observed with ensemble approach to check performance of IDS.

14.3 Methodology of Proposed System

In proposed IDS, an ensemble of supervised and unsupervised classifiers is done to improve performance. Classifiers used for ensembling are decision tree (DT), support vector machine (SVM) and XGboost. UNSW NB15 dataset is used for training all classifiers. Testing is performed in real-time environment to check robustness of IDS. To reduce training time of classifier, feature reduction techniques are used.

14.3.1 Signature Creation for Collaborative Attack Over UNSW NB15

UNSW NB15 dataset is generated in Australian Centre for cyber security (ACCS) in real-time environment consists of nine types of attack. Size of dataset is 100 GB with 49 features having last label for attack or normal data. The instances available for this dataset are two million and 540,044. The number of instances in the training set is 175,341, and the testing set is 82,332 instances [24]. This dataset provides improved performance over other dataset [25]. To generate signature of collaborative attack, scripts are written in scapy so that packets of attacks are created and sent in the network. Scapy allows creating packet for new attacks as per the features of dataset. Packets for some collaborative attacks are created which represents combination of multiple attacks. To generate attacks, probe and DoS attacks were combined together to create new attack. Signatures are created update with class label collaborative attack in dataset. Packets for collaborative attacks are captured in real-time environment where scripts are used to generate it. After merging all signatures to dataset, all classifiers are trained using dataset to test in real-time environment. Process for signature creation is represented in Fig. 14.1. Packets of collaborative attacks are randomly generated and penetrated into the network along with the normal traffic. Packets are captured using ethereal and saved in.pcap file. This file is passed for pre-processing and features are extracted from it.

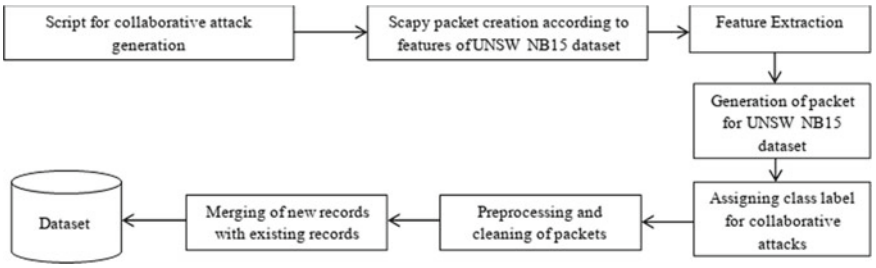


Fig. 14.1 Process for signature creation for collaborative attack

14.3.2 Feature Selection

Feature selection is used for reducing number of features from the dataset. It will help to reduce computational cost as well as training time for classifiers. Features should be reduced in manner that it will not affect performance or accuracy or detection rate of the IDS. Proposed IDS use information gain (IG) for feature selection which is used to calculate quality of each split done for finding the accuracy of the classifier. As decision tree, XGboost and SVM classifiers are used in proposed system, information gain provides better performance as compared to other techniques. Information gain values for each feature are calculated with formula given in Eq. 14.1. As shown in Eq. 14.1, information gain for all 49 features is calculated, and only features having good gain for information are selected. Number of features obtained after reducing it using information gain is 35 which will be used to train all classifiers used in proposed IDS.

$$\mathrm{IG} = -\sum_{i}^{c} p_i \log_2 pi \tag{14.1}$$

where c—number of classes for attacks.
pi—probability of random occurrence of selecting record of class.

14.3.3 Proposed System Architecture

As shown in Fig. 14.2, IDS is implemented using ensemble of DT, SVM and XGboost.

All classifiers were trained using reduced UNSW NB15 dataset with 35 features. As each classifier has its own limitation, it can happen that sometimes classifier results in biased that is wrong output. Ensemble technique is used to improve performance of the IDS. In ensemble, all the four classifier output is passed as input to the majority voting algorithm. This algorithm collects input from all four classifiers, and then according to the input, the final prediction is made. Final decision is made according

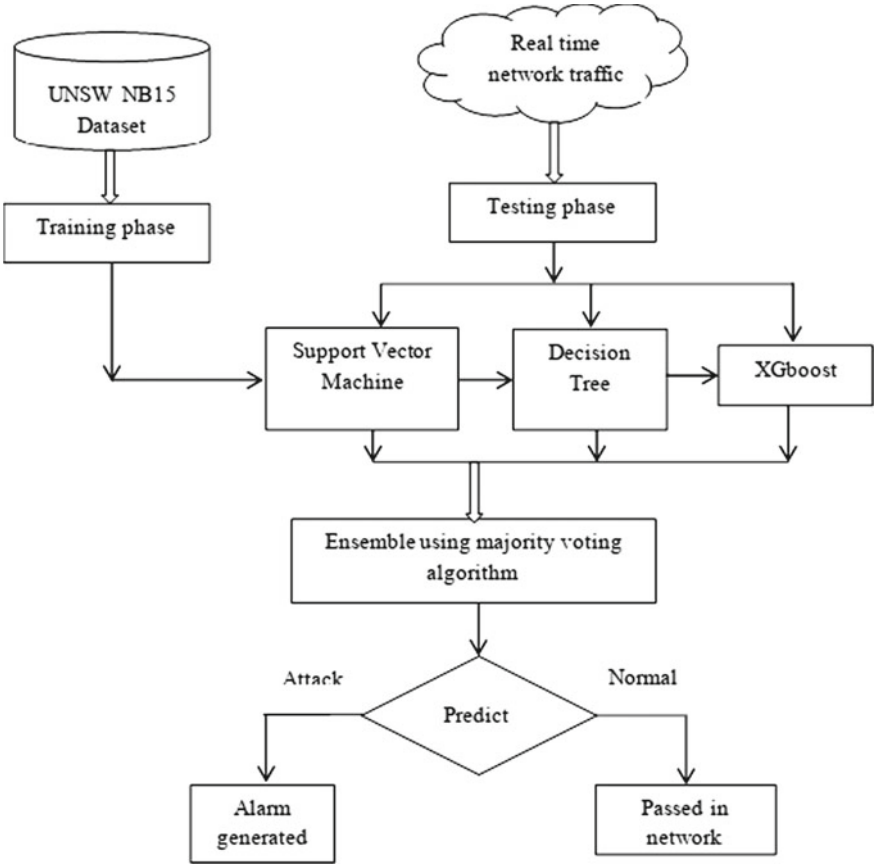


Fig. 14.2 Architecture of proposed ensemble-based IDS

to majority predictions received from all four classifiers. Using ensembling, biased and incorrect predictions can be reduced. If final prediction is attack, then alarm is generated for administrator. If prediction is normal, then packet is passed in the network. All classifiers performance along with ensembling is recorded in real-time environment with various performance parameters. IDS work efficiently and provides very good detection rate for both single and collaborative attack detection. This system provides reduced FAR and increases in accuracy.

14.4 Results and Discussions

Proposed IDS with ensemble of three classifiers is tested for single and collaborative attack detection with help of new updated dataset. Performance of the classifiers is

evaluated using precision, recall, accuracy and false alarm rate parameters. All classifiers are tested individually on UNSW NB15 dataset along with ensemble technique. Equations from 14.2 to 14.4 show Formula to calculate parameters precision, recall and accuracy.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14.2}$$

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{total}} \tag{14.3}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{14.4}$$

where TP—Prediction is positive value and the prediction is correct.
FP—Prediction is positive value, and the prediction is incorrect.
TN—Prediction is negative value, and the prediction is correct.
FN—Prediction is negative value, and the prediction is incorrect.

All these values are used to calculate precision, recall, false alarm rate and accuracy. Comparison of it is shown in Table 14.1.

Precision, recall and accuracy are used to show the robustness and completeness of classifier. Another parameter false positive rate is also used for evaluation. FAR is the total number of wrong predictions made out of actual value. All these parameters are tested for three classifiers individually and in ensemble approach. As per observations from Table 14.1, we can conclude that ensemble approach shows better performance in detection of single attack and collaborative attack. Graphical representation for all parameters with classifier is shown in Fig. 14.3. Figure shows reduction in false alarm rate for ensemble approach and improvement in precision, recall and accuracy of IDS.

As per shown in Fig. 14.3, we can observe that ensemble classifier provides better performance in precision, recall and accuracy as compared to individual classifier. Ensemble approach shows reduced FAR as compared to individual classifier. We can summarise that ensemble approach provides good exactness and completeness in terms of attack detection, increase of 2% accuracy value and 5% reduction in FAR. Ensemble approach removes prediction incorrectness and limitations of biased predication.

Table 14.1 Comparison of individual classifiers with ensemble approach for all parameters

Classifier	Precision	Recall	FAR	Accuracy (%)
DT	0.92	0.78	0.14	91.22
SVM	0.89	0.72	0.19	87.55
XGboost	0.98	0.75	0.15	95.79
Ensemble approach	0.99	0.85	0.07	97.96

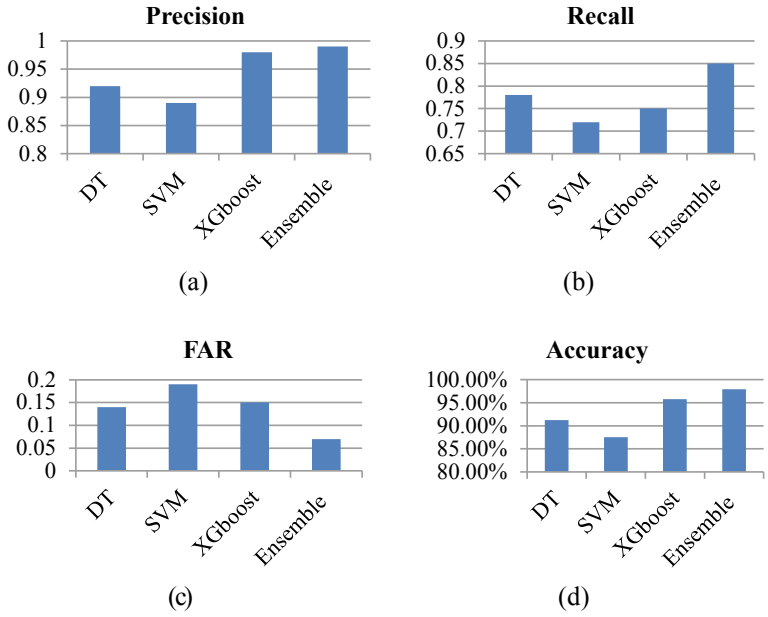


Fig. 14.3 **a**, **b**, **c**, **d** Graphical representation of comparison of individual classifier with ensemble approach for all parameters

14.5 Conclusion

An ensemble approach is presented in the paper which shows improved performance in terms of precision, recall, false alarm rate and accuracy. Most of the IDS are handling single attack at a time but this IDS will able to handle single as well as collaborative attacks. Signatures for some collaborative attacks are created and updated in existing dataset. Experiments are conducted on three classifiers as DT, SVM, and XGboost. All classifiers are trained using UNSW NB15 reduced 35 dataset features. To avoid biased and incorrect predictions, ensemble technique is used. Ensemble of classifier is done using majority voting algorithm which takes input from all three classifiers gives final prediction according to inputs. Experiments are conducted in real-time traffic, and ensemble approach shows better accuracy than single classifiers. Results prove the improvement of 2% in accuracy and reduction of more than 5% in false alarm rate. In future, proposed IDS can be modified for anomaly detection of unknown attacks. As proposed IDS use only signature-based detection, new or novel attacks cannot be identified and can harm a network. In future, signature and anomaly-based IDS can be combined together to form a new hybrid framework for attack detection.

References

1. Depren, O., Topallar, M., Anarim, E., Kemal, C.: An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Expert Syst Appl. **29**(4), 713–722 (2005)
2. Thakkar, A., Lohiya, R.: A review of the advancement in intrusion detection datasets. Proc. Comput. Sci. **167**, 636–645 (2020)
3. Saranya, T., Sridevi, S., Deisy, C., DucChung, T., AhamedKhan, M.: Performance analysis of machine learning algorithms in intrusion detection system: a review. Proc. Comput. Sci. **171**, 1251–1260 (2020)
4. Mebawondu, J., Alowolodu, O., Mebawondu, J., Adetunmbi, A.: Network intrusion detection system using supervised learning paradigm. Sci. African **9**, e00497 (2020)
5. Folino, G., Sabatino, P.: Ensemble based collaborative and distributed intrusion detection systems: aa survey. J. Netw. Comput. Appl. **66**, 1–16 (2016)
6. Tao, P., Sun, Z., Sun, Z.: An improved intrusion detection algorithm based on GA and SVM. IEEE Access **6**, 13624–13631 (2018)
7. Aburomman, A., Reaz, M.: A survey of intrusion detection systems based on ensemble and hybrid classifiers. Comput. Secur. **65**, 135–152 (2017)
8. Jabbar, M., Aluvalu, R., Satyanarayana, S.: RFAODE: a novel ensemble intrusion detection system. Proc. Comput. Sci. **115**, 226–234 (2017)
9. Divyasree, T.H., Sherly, K.K.: A network using intrusion detection system based approach on ensemble cvm efficient feature selection using efficient feature selection approach. Proc. Comput. Sci. **143**, 442–449 (2018)
10. Vinutha, H.P, Poornima, B.: An ensemble classifier approach on different feature selection methods for intrusion detection, Inf. Syst. Des. Intell. Appl. **672**, 442–451 (2018)
11. Gupta, G., Kulariya, M.A.: Framework for fast and efficient cyber security network intrusion detection using apache spark. Proc. Comput. Sci. **93**, 824–831 (2016)
12. Saranya, T., Sridevi, S., Deisyc, C., Chungd, T., Khane, M.: Performance analysis of machine learning algorithms in intrusion detection system: a review. Proc. Comput. Sci. **171**, 1251–1260 (2020)
13. Kunal, Dua, M.: Attribute selection and ensemble classifier based novel approach to intrusion detection system. Proc. Comput. Sci. **167**, 2191–2199 (2020)
14. Gaikwad, D.P., Thool, R.: Intrusion detection system using bagging with partial decision tree base classifier. Proc. Comput. Sci. **49**, 92–98 (2015)
15. Almasoudy, F., Al-Yaseen, W., Idrees, A.: Differential evolution wrapper feature selection for intrusion detection system. Proced. Comput. Sci. **167**, 1230–1239 (2020)
16. Niksefat, S., Kaghazgaran, P., Sadeghiyan, B.: Privacy issues in intrusion detection systems: a taxonomy survey and future directions. Comput. Sci. Rev. **25**, 69–78 (2017)
17. Khonde, S.R., Ulagamuthalvi, V.: Hybrid framework for intrusion detection system using ensemble approach. Int. J. Adv. Trends Comput. Sci. Eng. **9**(4), 4881–4890 (2020)
18. Gu, J., Wang, L., Wang, H., Wang, S.: A novel approach to intrusion detection using SVM ensemble with feature augmentation. Comput. Secur. **86**, 53–62 (2019)
19. Aljawarneh, S., Aldwairi, M., Yassein, M.: Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. J. Comput. Sci. **25**, 152–160 (2018)
20. Hubballi, N., Suryanarayanan, V.: False alarm minimization techniques in signature-based intrusion detection systems: a survey. Comput. Commun. **49**, 1–17 (2014)
21. Li, Y., Wang, J., Tian, Z., Lu, T., Young, C.: Building lightweight intrusion detection system using wrapper-based feature selection mechanisms. Comput. Secur. **28**(6), 466–475 (2009)
22. Ring, M., Wunderlich, S., Scheuring, D., Landes, D., Hotho, A.: A survey of network-based intrusion detection data sets. Comput. Secur. **86**, 147–167 (2019)
23. Moustafa, N., Slay, J.: UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: Military Communications and Information Systems Conference (MilCIS), IEEE Explore (2015)

24. Moustafa, N., Slay, J.: The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset. Info. Secur. J.: A Global Persp. 1–14 (2016)
25. Sonule, A., Kalla, M., Jain, A., Chouhan, D.: Unsw-Nb15 dataset and machine learning based intrusion detection systems. Int. J. Eng. Adv. Technol. (IJEAT) **9**(3), 2638–2648 (2020)

Chapter 15
A Decision Based Asymmetrically Trimmed Modified Geometric Mean Algorithm for the Removal of High Density Salt and Pepper Noise in Images and Videos

K. Vasanth, C. N. Ravi, S. Nagaraj, M. Vadivel, K. Gopi, S. Pradeep Kumar Reddy, and Thulasi Prasad

Abstract A decision based asymmetrically trimmed modified geometric mean algorithm (DBATMGMA) for the removal of high density salt and pepper noise in images and videos is proposed. The algorithm uses fixed 3×3 windows for increasing noise densities. The algorithm initially checks for the presence of outliers (0 or 255) in the processed pixel. If it holds the outliers, the processed pixel is termed faulty. If the processed pixel is faulty, check for the four neighbors are noisy or not (0 or 255). If all the four neighbors are noisy, then check for all the pixels is noisy in the confined neighborhood. If the entire window is noisy, then the corrupted pixel is replaced by mean of all the elements else the corrupted pixel is replaced by mean of the four neighbors. If all the four neighbors are not noisy, then the faulty pixels are replaced with asymmetrically trimmed modified geometric mean. If the pixels does not hold the outlier, then the pixel is considered undamaged and left unaltered. The proposed algorithm is compared with standard algorithms on an image database. The noise suppression along with information preservation capability of the proposed algorithm was found to be very good both in terms of qualitative and quantitative measures.

K. Vasanth (✉) · M. Vadivel · S. P. K. Reddy · T. Prasad
Department of ECE, Vidya Jyothi Institute of Technology, Hyderabad, Telangana, India

C. N. Ravi
Department of EEE, Vidya Jyothi Institute of Technology, Hyderabad, Telangana, India

S. Nagaraj
Department of ECE, Sri Venkateswara College of Engineering and Technology, Chitoor, Andhra pradesh, India

K. Gopi
Department of ECE, Sreenivasa Institute of Technology and Management Studies, Chitoor, Andhra pradesh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_15

15.1 Introduction

Fixed valued impulse noise also called as salt and pepper noise (SPN) distorts image due to transmission errors. Fixed valued impulse noise with equal noise probability for an 8 bit image for very pixel is given by the noise model in Eq. 15.1. If $K\ (i, j)$ takes pixel values between 0 to 255 for all (i, j) for an 8 bit image, then noise model is given as

$$O(i,\ j) = \begin{array}{l} 0 \text{ for probability } p \\ K(i,\ j) \text{ for probability } 1 - p - q \\ 255 \text{ for probability } q \end{array} \tag{15.1}$$

where, p+q gives the noise level

A nonlinear operation such as median filter is best suited to eliminate fixed valued impulse noise [1]. Over the years, many nonlinear filters have been proposed and few are listed here below. An adaptive median filter is good in elimination of corrupted pixels but a large window size blurs the image and a smaller window lacks information to restore the image [2]. A decision-based median filter (DBA) [3] used median only for corrupted pixels. At high noise densities a pre-processed neighbor. This induced streaking of edges in images. An asymmetrically trimmed decision-based median filter [4] was proposed. This gave very good noise elimination capability but induced fading effect at very high noise densities. This fading was controlled by an improved algorithm [5] which used global mean instead of local mean if all nine pixels of processing window are corrupted. A new class of trimmed modified Winsorized filter was proposed. These algorithms used modified Winsorized mean [6] or median [7] or variants [8] for the removal of high density salt and pepper noise. All the algorithms work well in eliminating the high density salt and pepper noise. But these algorithms induce ambiguity while eliminating the high density noise. An algorithm has to be formulated for the elimination of high density salt and pepper noise without inducing ambiguity. The attempt in this paper is to achieve the above is by keeping a fixed 3×3 window. The proposed algorithm will eliminate the noise proposed in Eq. 15.1. Section 15.2 gives the flowchart of the proposed algorithm. Section 15.3 gives the simulation results and discussions. Section 15.4 gives the conclusion of the work.

15.2 Decision-Based Neighborhood Referred Unsymmetrical Trimmed Modified Geometric Mean (DBNRUTMGMA)

The decision-based neighborhood referred asymmetrical trimmed modified geometric mean algorithm (DBNRATMGMF) is introduced for the removal of high

density salt and pepper noise. The DBNRATMGMF filter will also be referred as proposed algorithm. The conventional geometric mean is not stable for increasing noise densities. In order to make it stable, few additional conditions were included as part of the algorithm. Hence, the instability found in conventional geometric mean is eluded in modified geometric mean. In case of videos, the frames are converted into images, and the proposed algorithm is applied on each image and appended later to for a restored video. The flowchart of the proposed algorithm is given in Fig. 15.1.

15.3 Simulation Results

The images used in the paper was used from website of USC Viterbi School of Engineering SIPI database [9]. The noise was added on images from 10 to 90% in increments of 10%. Different algorithms were tested on the noisy images, and performance were tabulated using the quantitative measures such as peak signal to noise ratio (PSNR) and structural similarity index metric (SSIM). The noise removal capability of the algorithm is calculated using the metrics given in Eqs. 15.2, 15.3, and 15.4, respectively.

$$\text{PSNR} = 10 \log 10 \left(\frac{255^2}{\text{MSE}} \right) \tag{15.2}$$

$$\text{MSE} = \frac{\sum_i \sum_j (I_{ij} - R_{ij})2}{\text{size of processed Image}} \tag{15.3}$$

I indicates reference image, and R indicates the filtered image. The SSIM metric is evaluated on image using many 3×3 windows maneuvered in an image. In any area of the image arbitrarily, the evaluation of common window size between two equal sized windows is given by Eq. 15.3.

$$\text{SSIM}(x, y) = \frac{(2\mu_I \mu_R + D1)(2\sigma_I \sigma_R + D2)}{(\mu_I^2 + \mu_R^2 + D1)(\sigma_I^2 + \sigma_R^2 + D2)} \tag{15.4}$$

where μ refers to average, and σ indicates the standard deviation of small windows. [10]. The value of the PSNR should lie between 20–24db for subsequent processing of images. A value closer to 1 indicates the structural preservation property which is close to the original image. To quantify the good result of the algorithms additional metrics such as image enhancement factor (IEF), Pratt's figure of merit (Pratt's FOM), normalized cross-correlation (NCC), blur metric (BM) and universal quality index (UQI). Tables 15.1 and 15.2 give the peak signal to noise ratio (PSNR) and structural similarity index metric (SSIM) performance of different algorithms on bridge image corrupted by fixed valued impulse noise. It was inferred from Table 15.1 that high value of PSNR indicates very good noise elimination capabilities of the proposed

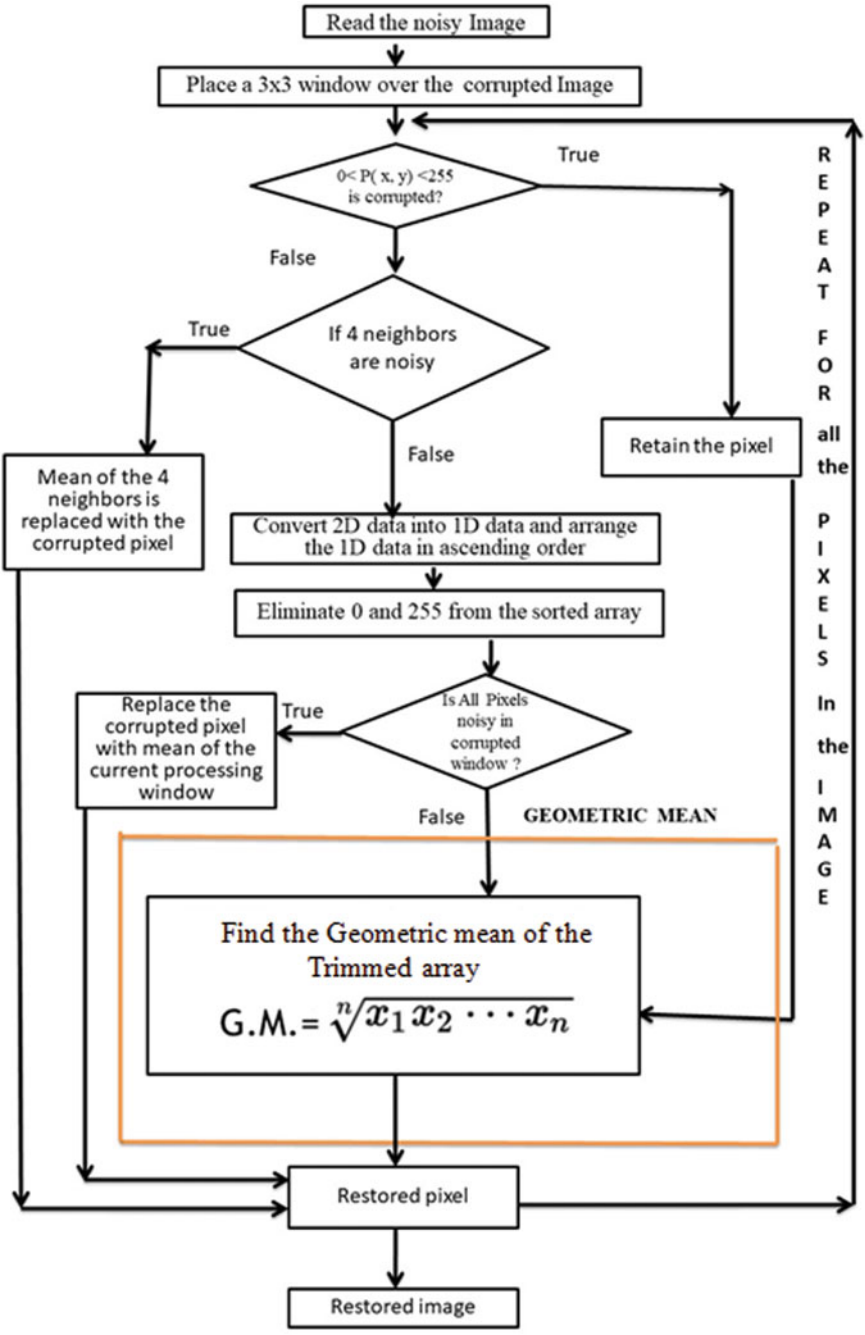


Fig. 15.1 Flowchart of decision-based neighborhood referred asymmetrical trimmed modified geometric mean

Table 15.1 Measure of PSNR of various algorithms on bridge image corrupted by SPN

ND	DBA [3]	MDBUTMF [4]	MDBUTMF_GM [5]	Win mean [6]	Win median [7]	PA
10	28.75	34.68	34.67	34.31	34.44	34.68
20	28.04	31.28	31.21	31.31	31.02	31.61
30	26.56	29.04	28.98	29.33	28.88	29.46
40	25.24	27.16	27.27	27.78	26.98	27.83
50	24.18	25.53	26.68	26.52	25.49	26.51
60	23.08	23.98	23.68	25.17	24.05	25.25
70	21.77	22.23	22.14	23.83	22.20	23.68
80	20.37	20.72	20.34	22.20	20.13	22.06
90	NAN	18.42	18.34	20.04	NAN	19.52

Table 15.2 Measure of SSIM of various algorithms on bridge image corrupted by SPN

ND	DBA [3]	MDBUTMF [4]	MDBUTMF_GM [5]	Win mean [6]	Win median [7]	PA
10	0.928	0.978	0.979	0.977	0.978	0.979
20	0.911	0.955	0.935	0.954	0.953	0.956
30	0.879	0.924	0.925	0.927	0.923	0.929
40	0.836	0.884	0.885	0.895	0.882	0.896
50	0.784	0.832	0.834	0.855	0.831	0.857
60	0.720	0.763	0.763	0.805	0.764	0.805
70	0.644	0.666	0.662	0.737	0.664	0.733
80	0.543	0.529	0.529	0.638	0.523	0.628
90	0	0.342	0.340	0.484	0	0.458

algorithm. It was also noted that structure of the image was well preserved even at very high noise densities. Table 15.3 illustrates the performance of proposed algorithm on different images of the database corrupted by 90% SPN for different image measures. In order to evaluate the information preservation performance of the algorithms, a synthetic image was created using eight visually differentiable gray levels. The image was corrupted by fixed valued impulse noise, and the algorithms were applied on it. It was vivid from image 2 that the proposed algorithm preserved ramp edge but attenuates the roof and step edge, respectively. Ramp edge corresponds to different gray levels of the image. The attenuated similar look alike gray levels are referred to as step image, and they are attenuated by the proposed algorithm. In case of videos as shown in Fig. 15.3, the performance of proposed algorithm on different videos such as viptraffic.avi, rhino.avi, shuttle.avi is corrupted by 50% SPN. It was found that proposed algorithm is very good in noise suppression and information preservation. Hence, this algorithm will be suitable for low-level processing of images corrupted by high density salt and pepper noise. The good performance of the proposed algorithm

Table 15.3 Measures of the proposed filter on various database images with 90% of SPN

Images	PSNR	IEF	MSE	SSIM	FOM	Blur metric	NCC	UQI
aero.tiff	30.43	396	58	0.729	0.352	0.321	0.982	0.255
barbara.tif	21.31	35	480	0.600	0.449	0.366	0.958	0.413
bird.tif	24.61	73	224	0.812	0.480	0.432	0.991	0.428
elaine.png	23.86	62	267	0.645	0.490	0.412	0.982	0.435
fingerprint.bmp	17.15	13	1252	0.443	0.506	0.310	0.919	0.440
Hill.jpg	23.75	62	272	0.594	0.422	0.352	0.985	0.397
Butterfly.bmp	19.92	25	661	0.656	0.476	0.375	0.923	0.396
peppers.gif	18.92	20	833	0.663	0.455	0.425	0.945	0.496
two.bmp	18.25	16	971	0.623	0.467	0.327	0.926	0.390
Pentogon.tiff	23.81	56	269	0.563	0.468	0.334	0.981	0.355
Zelda.png	26.35	114	150	0.768	0.399	0.438	0.979	0.563
House.bmp	22.93	50	331	0.713	0.444	0.366	0.986	0.331
Crater.png	26.48	104	146	0.650	0.445	0.367	0.989	0.358
Cameraman.bmp	16.35	12	1503	0.580	0.448	0.258	0.981	0.272

is mainly due the fact that the asymmetrical trimming of outliers from the data set for which geometric mean is evaluated. The values computed do not use outliers. Hence, the computed modified geometric mean is outlier free. Hence, the performance of the algorithm was found good at very high noise densities (Figs. 15.2, 15.3 and 15.4).

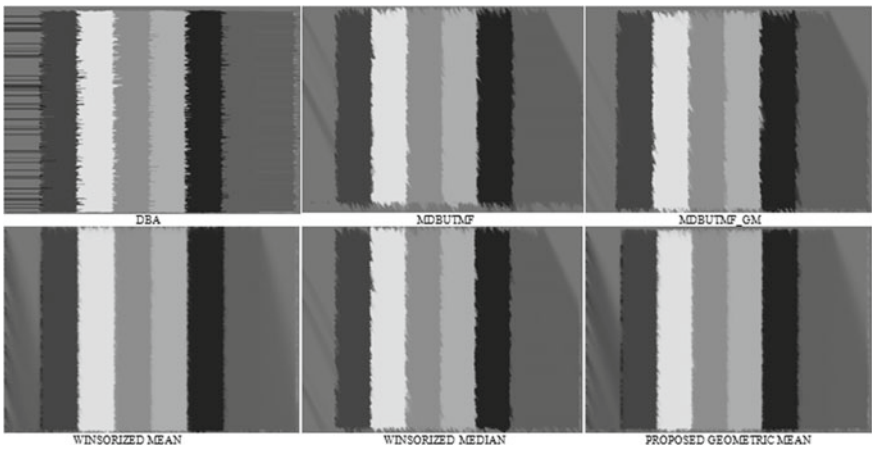


Fig. 15.2 Visual results of different filters/algorithms on ramp synthetic image



Fig. 15.3 Visual results of proposed filter on various images of the database corrupted by 90% SPN

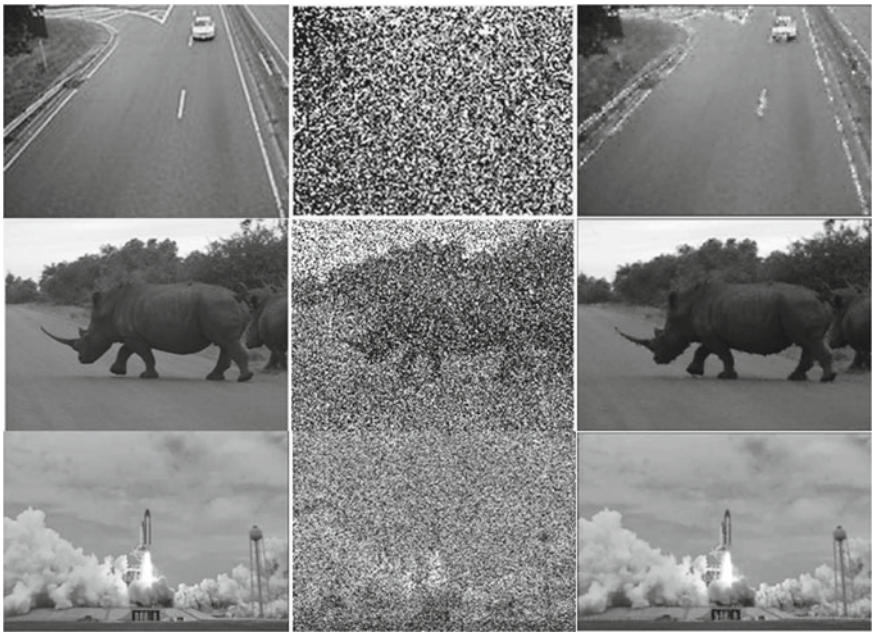


Fig. 15.4 Visual results of decision-based neighborhood referred asymmetrical trimmed modified geometric mean on videos (viptraffic, rhinos, shuttle) corrupted by 50% SPN

15.4 Conclusion

A decision-based neighborhood referred asymmetrically trimmed modified geometric mean filter is proposed for the removal of high density salt and pepper noise in images and videos. The proposed algorithm used a fixed 3×3 kernel at high noise densities. The proposed algorithm does not induce any artifacts during the

noise removal. Hence, this algorithm was a suitable for high density SPN removal for both images and videos.

References

1. Astola, J., Kusmanen, P.: In: Fundamentals of Non Linear Digital Filtering, CRC press (1997)
2. Hwang, H., Haddad, R.A.: Adaptive median filters: new algorithms and results. IEEE Trans. Image Process. **4**, 499–502 (1995)
3. Srinivasan, K.S., Ebenezer, D.: A new fast and efficient decision-based algorithm for removal of high-density Impulse noises. IEEE Signal Process. Lett. **14**(3), 189–192 (2007)
4. Esakkirajan, S., Veerakumar, T., Subramanyam, A.N., Prem Chand, C.H. : Removal of high density Salt and pepper noise through modified decision based unsymmetrical trimmed median filter. IEEE Signal Process. Lett. **18**(5), 287–290 (2011)
5. Veerakumar, T., Esakkirajan, S., Vennila, I.: An approach to minimize very high density salt and pepper noise through trimmed global mean. Int. J. Comput. Appl. **39**–12, 29–33 (2012)
6. Vasanth, K., Manjunath, T.G., NirmalRaj, S.: A decision based unsymmetrical trimmed modified winsorized mean filter for the removal of high density salt and pepper noise in images and videos. Proc. Comput. Sci. **54**, 595–604 (2015)
7. Christo, M.S., Vasanth, K., Varatharajan, R.: A decision based asymmetrically trimmed modified winsorized median filter for the removal of salt and pepper noise in images and videos. Multimed. Tools Appl. **79**, 415–432 (2020)
8. Vasanth, K., Varatharajan, R.: An adaptive content based closer proximity pixel replacement algorithm for high density salt and pepper noise removal in images. J. Ambient Intell. Human Comput. (2020). https://doi.org/10.1007/s12652-020-02376-2
9. Signal and image processing institute, University of South California https://sipi.usc.edu/database/ Accessed on 10 March (2020)
10. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error measurement to structural similarity. IEEE Trans. Image Process. **13**(4), 102–109 (2004)

Chapter 16
Anti-poaching System Using Wireless Sensor Networking



T. Asha, H. S. Jahanavi, Sony S. Rao, S. Sinchana, and K. A. Priyanka

Abstract Our life is dependent upon trees. There is a deep relationship between man and trees. Nowadays, there are many incidents about smuggling of trees like sandal and sagwan. These trees are very costly and meager. They are used in the medical sciences, cosmetics and many more. To restrict their smuggling and to save forests around the globe, some preventive measures need to be deployed. A system has been proposed that uses RFID tags and four sensors along with Wi-Fi module which can be used to restrict smuggling. The proposed system is cost effective and efficiently signals the forest officer or any authorities regarding any illegal smuggling activities.

16.1 Introduction

Forests represent more or less half hour of the worldwide acreage. They provide environment for both humans and a few species which share the product of the precious ecosystem. Amerciable research represents one amongst the largest challenges of forests property. Preparation is not associated exclusively with Bharat; China, Australia and African countries are also fighting the same issue. Indian sandalwood prices 12,000 to 13,000 agencies per kilogram while red sandalwood prices INR 10 crore per ton in the international market. The Indian tree has become scarce in recent years. The Indian Government is seeking to restrict the exportation of wood in an attempt to control its attainable loss. Forests provide clean water and air, wood product timber, living environments, fertile soil and leisure opportunities, and these trees make the environment beautified. In addition, they are an important economic resource, which produces highly marketable timber. Despite this amount, ample forestland is being eliminated from forest development, largely due to urban area pretending to be wooded areas, which will remain replaceable.

Wood smuggling has generated problems of social, cultural, legal, law and order in areas bordering Bharat. Trees, which are primarily influenced by the square measure,

T. Asha (✉) · H. S. Jahanavi · S. S. Rao · S. Sinchana · K. A. Priyanka
Department of ECE, GM Institute of Technology, Davangere, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_16

represent timber, teak timber and rosewood. There are several measures undertaken by entirely different organizations, and particularly Bharat's Government, to ease this drawback. It involves the housing, coaching and forest-wide training of anti-poaching seekers. Nevertheless, some of the precautionary steps remain ineffective. The highest hopeful resolution is—"Protecting valuable trees from smuggling using RFID and sensors" which can be a strong, reliable and practical technology for watching. In addition, a machine-controlled smart device is developed to fix these issues.

These modules are supplied by the combination of the new wireless communication systems and embedded solutions. Each tree should have one very small system unit embedded with: sensors and RFID tags. The distance of elements higher than the aforementioned will submit this tree status to the server, using Wi-Fi module. The information is sent via twitter, and IoT organize is formed here from now on. The identified data is deciphered by the BLYNK APP, which maintains the information of each of those trees. In cloud, information is investigated to see whether the tree is being covered or pushed down. Nevertheless, the information continues to be constantly informative about the condition of the trees, and the forest specialists should discuss the circumstances of the crisis. Section 16.1 briefly explains about the introduction, problem statement and objective of the project. Chapter 16.2 briefly explains the different papers referred for the literature survey. Sections 16.3 and 16.4 give the brief description about the system design and implementation. Results and conclusion are discussed in the Sects. 16.5 and 16.6.

16.2 Literature Survey

In this section, some of the relevant papers have been studied to gain insight into the existing systems.

Raghavendra et al. [1] proposed IoT-based anti-smuggling system for trees in forest where the main objective of the system is to restrict the smuggling and save the valuable trees so to maintain balanced ecosystem by preventing deforestation. The system uses the GPS technology from which they can find the location of the tree where the poaching is done. It also uses a chip (board) on which various sensors are embedded which are controlled through IoT. These sensors supervise and determine the parameters like tilting, burning and cutting of the trees, and these are accessed on android app installed in android smartphone. Pushpalatha et al. [2] proposed an IOT-based GSM-based real-time forest anti-smuggling monitoring system where the proposed work foiled the plan of an enemy of pirating framework which is valuable in protected forest regions. The thieving and illegal business trees production, such as sandalwood, teak, sagwan and so on, is an important concern. It is a more burglary to flora and fauna of the forests. As residence involved, this belief system is built to combat such sneaking exercise by using the most recent advances.

Naveenraj et al. proposed an IoT-based anti-poaching alarm system for trees in forest to prevent sneaking [3]. The structural frame uses three sensor tilt sensors to

recognize the tree's propensity once it is cut, temperature-sensing element to identify timberland fires and sound sensing elements to successfully discover illegal logging. Information obtained from those sensors is calculated consistently. Ringer is enacted for the tilt sensing and sound sensing function, and a water syphon is actuated for temperature sensing portion. Using the Wi-Fi module, generated information is stored in the cloud server. When any occasion arises, forest authorities are suggested, so that correct step can be made. Suguvanam et al [4] presented an idea for innovative protection of valuable trees from smuggling using RFID and sensors. This paper presents a pattern of preventing the importation into the forest of valuable trees [4] such as sandalwood tress, red sandal trees by using GSM module and RFID. In the proposed device, vibration sensors and continuity sensor are used to find scaling down of valuable trees instantly. Here, the vibration sensor is used to detect the disturbance produced when the trees are being cut and the continuity sensor is used to test the continuity between the neighboring trees. RFID is used within the current device to find the missing tags in the tree.

Kotkar et al [5] put forward an anti-smuggling system for trees in forest using flex sensor and Zigbee. They proposed a system which able to restrict smuggling of tree in the forest system using GSM module, Zigbee module, microcontroller and flex sensor. The program also essentially aims to reduce global warming. The species natural environment is also partially conserved. Through this device the trees can also be covered from forest fire. It is in effect tends to reduce the deforestation. Sonwane et al. [6] proposed microcontroller-based design and development of wireless sensor node for anti-poaching capable of exposing theft by keeping track of vibrations produced during the cutting trees using a MEMS accelerometer.

16.3 System Design

The block diagram of anti-poaching system which consists of ESP8266, flame sensor, metal detector sensor, vibration sensor and RFID is shown in Fig. 16.1. The main idea is to design a transportable wireless detector node, which will be a part of wireless detector network. This method includes two modules: One is the sensor module, and the other is the controller module which will be at the tree spot. The information can be received from the detector continuously in the cloud. This is an IOT-based project where information about the detector is continuously uploaded to the cloud server over a Wi-Fi module.

In case of temperature or fire sensor, buzzer is turned on at the time of forest fire. Fire sensor is used to detect the flame or fire. When fire burns, it emits a small amount of infrared light; this light will be received by the photodiode on the sensor module. Fire detector consists of photodiode, resistor, capacitor, potentiometer and LM393 comparator in an IC.

In case of metal sensor, if any metal like axe is came in its range, it will sense and signals are sent to cloud. The operation of metal sensor is based upon the principles of electromagnetic induction. Metal detectors contain one or more inductor coils

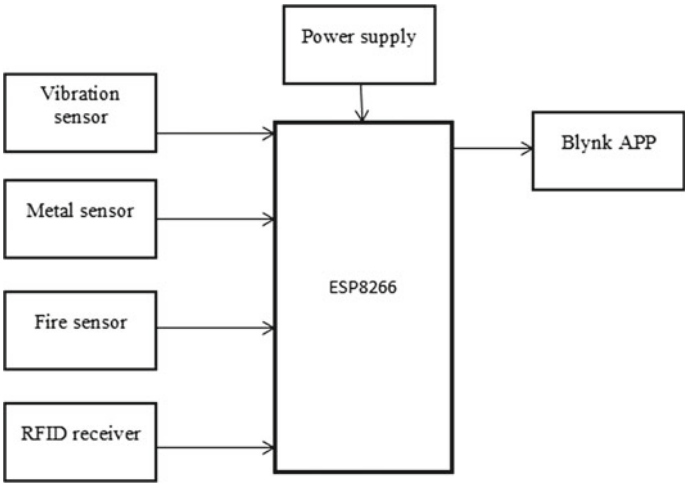


Fig. 16.1 Block diagram of anti-poaching system using wireless sensor networking

that are used to interact with metallic elements. A pulsing current is applied to the coil, which then induces a magnetic field. When the magnetic field of the coil moves across metal, the field induces electric currents (called eddy currents). The eddy currents induce their own magnetic field, which generates an opposite current in the coil, which induces a signal indicating the presence of metal.

RFID is an identification tool for frequency. A simple, low-cost and automatic identification technology uses RF frequency to transfer information between an RFID reader and an RFID tag. It carries a small amount of distinctive information—a serial range or specific characteristic of the object. The data is scanned from a distance, no contact or maybe line of sight required. RFID technology uses an RFID tag, which is a small item, like a paste sticker, which can be hooked up to or inserted into a device. RFID tags include antennas to adjust them from an RFID transceiver to receive and respond to radio-frequency queries.

The vibration sensor type used is SW-18015P/20p, where series are unit spring form, no directional vibration detector trigger switch, and no angle can activate. The switch is an OFF-state electrical circuit, as it is still, when an external force comes into contact and equal vibration, or movement speed provides centrifugal (partial) power, semiconducting system feet can generate instant physical phenomenon is ON-state instant, once the external force is gone, turn back to OFF-state electric circuit. Switch has the identification P at the bottom of the rock to be completely airtight; the lifetime of two lakhs cycles can be reached within the conventional switch period. This series switches on high sensitivity control.

16.4 System Implementation

In these section, implementation details of system design in previous section are explained. Figure 16.2 shows the circuit diagram of anti-poaching system.

16.4.1 Flame Sensor

Flame sensor is used to detect fire flames. The fire sensor module consists of fire sensor and comparator to detect fire up to a range of 1 m. The flame sensor consists of D0 pin, ground pin and VCC pin. D0 pin of flame sensor is connected to the D2 pin of NodeMCU ESP8266, GND pin of flame sensor is connected to GND pin of ESP8266, and VCC pin is connected to 3V3 pin of ESP8266. Input to this module is flame or smoke, which is detected while burning. When the flame is detected, the system sends notification through Blynk application that fire is detected.

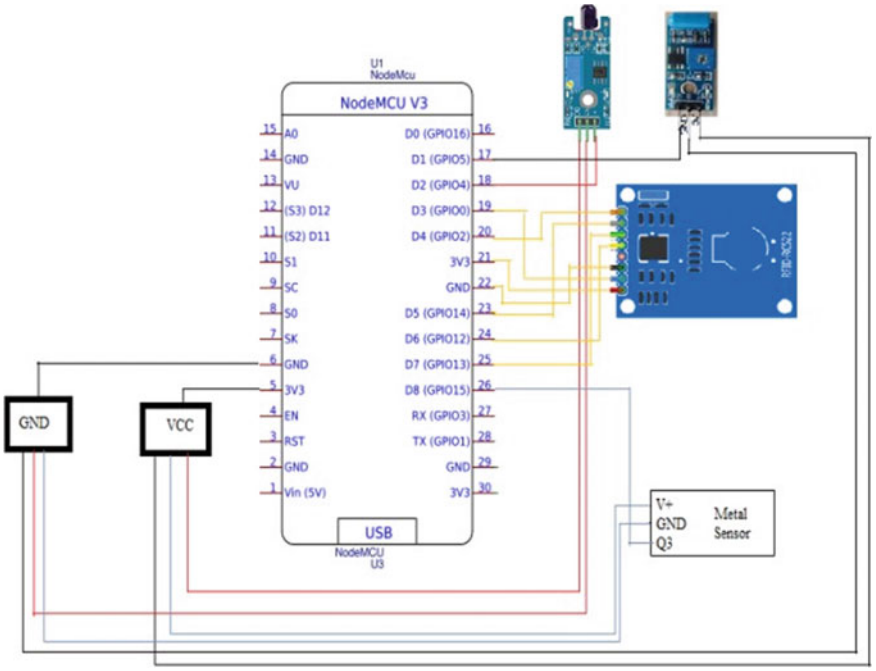


Fig. 16.2 Circuit diagram of anti-poaching system using wireless sensor networking

16.4.2 Vibration Sensor

The vibration-sensing element is also called as piezoelectric sensor. This sensor is used to measure the intensity of vibration. Input to this system is the vibration which is detected while a tree is being cut. When this sensor detects continuous vibration, a notification will be sent to the android phone through Blynk application showing the degree of intensity of vibration. This sensor consists of three pins, namely D0 pin, GND pin and 5 V pin. D0 pin of vibration sensor is connected to the general-purpose D1 pin of ESP8266, GND pin is connected to GND pin, and 5 V pin is connected to 3V3 pin of ESP8266.

16.4.3 Metal Detector Sensor

Metal detector sensor is used to detect the presence of metal. The metal detector sensor module operates by the currents caused in metal objects and responding once it happens. The buzzer rings when the metal is detected. The metal detector sensor consists of three pins, namely D0 pin, GND pin and VCC pin. D0 pin of the metal detector sensor is connected to the D8 pin of ESP8266, GND pin to GND pin and VCC to 3V3 pin of ESP8266. Input to this sensor is the presence of metal. When metal is detected by the sensor, a notification will be sent to the user through Blynk application that metal is detected.

16.4.4 Radio-Frequency Identification (RFID)

Radio-frequency identification (RFID) uses magnetic fields to mechanically establish and track tags connected to things. RFID is one methodology of automatic identification and data capture. The tags contain electronically hold on data. Input to this module is the presence of RFID tags or card. Trees will be provided with an RFID tag. When a tree is cut, RFID tag will be lost and notification will be sent to the user that a tag is missing. RC522 sensor which is a RFID module consists of VCC, reset, GND, IRQ, MISO/SCL/TX, MOSI, SCK and SS/SDA/Rx. RST pin of RC522 sensor is connected to the D3 pin ESP8266, SDA(SS) pin is connected to D4 pin of ESP8266, SCK pin is connected to D5 pin, MISO(SCL) pin is connected to D6 pin, and MOSI pin of RC522 pin is connected to the D7 pin of ESP8266.

16.4.5 Bylnk App

Blynk could be a new platform allowing to quickly create interfaces for managing and track IOS and android hardware ventures. Afterward, a project dashboard is created by installing the Blynk software and organize buttons, sliders, diagrams and other on-screen widgets. Pins can be switched on using the widgets and turn off or view sensor data. Whatever the project is, probably there are many tutorials that render the hardware component relatively straightforward, but software interface construction is very tough. However, the software side is much simpler than the hardware, with Blynk. Blynk is suitable for interfacing with basic tasks such as the temperature control of tank or remotely switch lights on and off. Blynk actually supports most of the Arduino modules, versions Raspberry Pi, ESP8266, ESP32, particle core and a couple of microcontrollers and single-panel computers. Wi-Fi and Ethernet shields are powered by Arduino but can also monitor devices that are often attached to the USB port of a device.

16.4.6 Arduino Ide

Arduino has integrated IDE development environment, a cross-platform application written in the Java programming language. A program for Arduino which is written with the IDE is called a sketch. Sketches are stored on the production machine as text files with the.info file extension, pre-1.0 sketches stored by Arduino software (lDE) with the.pde extension. The Arduino IDE supports the languages C and C++ using special rules of code structuring.

16.5 Results

The snapshots of different steps involved in designing and development of anti-poaching system using wireless sensor networking are shown Figs. 16.3, 16.4, 16.5 and 16.6.

16.6 Conclusion

The theory was completed in order to abstain from bringing valuable trees in protected woodland region. There are various approaches to securing trees, and here, an excellent technique has been done to connect a few sensors around trees with an IoT system. The idea was to recognize each tree as a smart tree (with a controller device and sensors) and to put several such trees under a network (Internet of Things). This

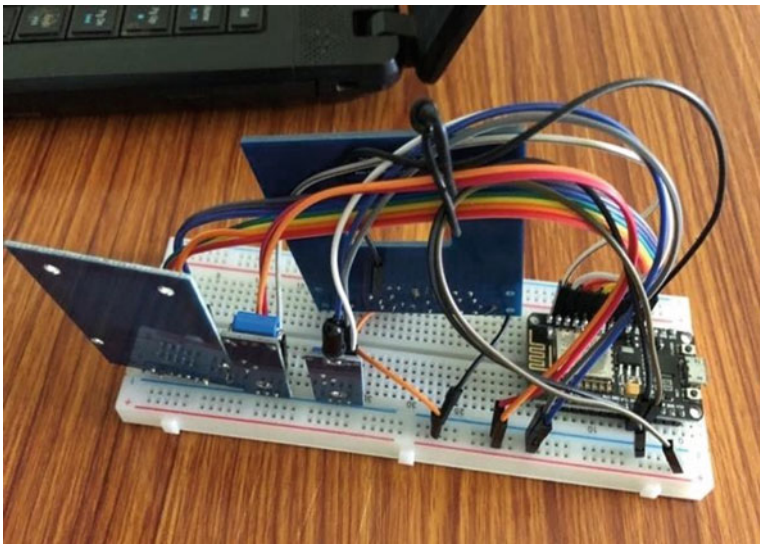


Fig. 16.3 Snap shot of Anti-poaching system

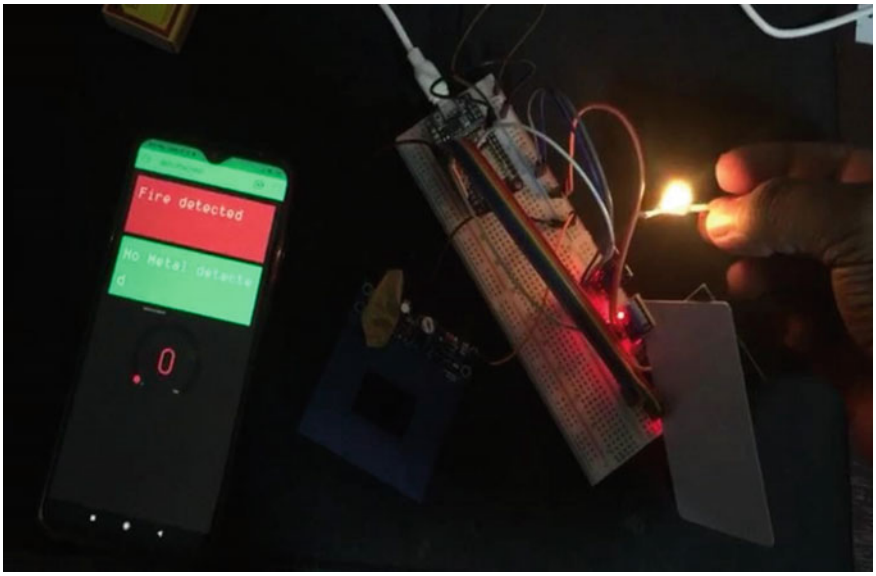


Fig. 16.4 Snapshot of metal detected shown in Blynk app

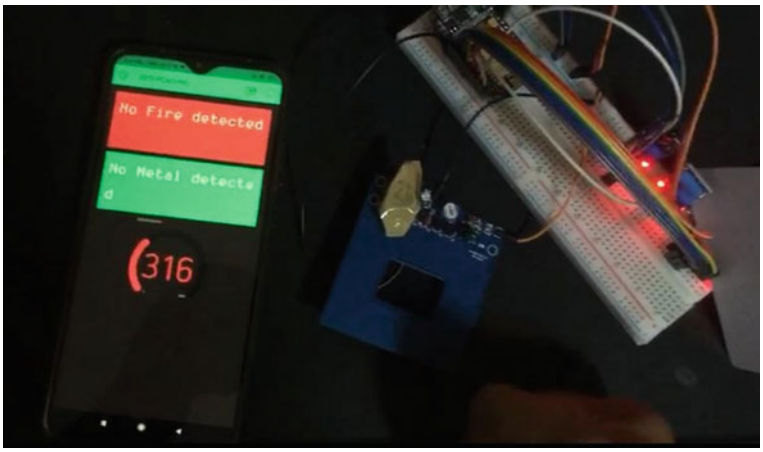


Fig. 16.5 Snapshot of vibration detected shown in Blynk app

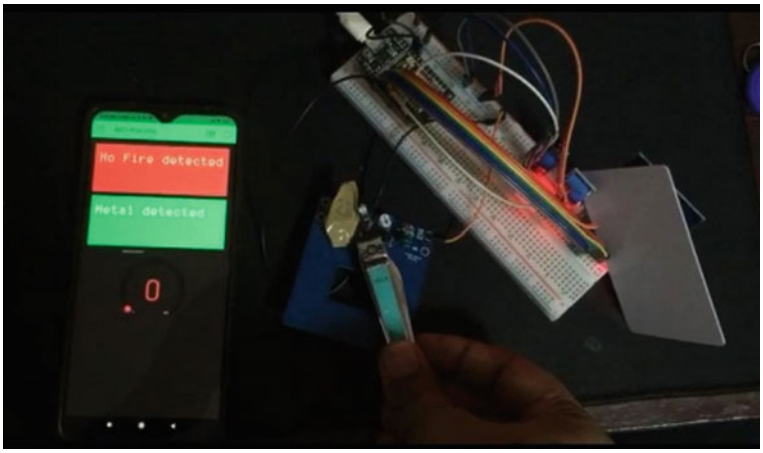


Fig. 16.6 Snapshot of fire detected shown in the Blynk app

system is designed to limit the sneaking of trees in the backwoods where individuals are not present to provide security. Such a system is designed in the backwoods where trees are exorbitant and where their protection is fundamental reality.

References

1. Hegde, R.I.: IoT Based anti-smuggling system for trees in forest. Int. J. Sci. Res. Rev. **8**(6), (2019)

2. Pushpalatha, R., Darshini, M.S.: Real time forest anti-smuggling monitoring system based on IOT using GSM. Int. J. Adv. Res. Comput. Commun. Eng. **8**(2), (2019)
3. Naveenraj, M., Arunprasath, Jeevabarathi, C.T., Srinivasan, R.: IoT Based anti-poaching alarm system for trees in forest. Int. J. Innov. Technol. Explor. Eng. (IJITEE) **8** (6S), (2019). ISSN: 2278–3075
4. Suguvanam, K.R., Senthil Kumar, R., Partha Sarathy, S., Karthick, K., Raj Kumar, S.: Innovative protection of valuable trees from smuggling using rfid and sensors. JIRSET (2017)
5. Kotkar, N.R.: Anti smuggling system for trees in forest using flex sensor and zigbee. IJARCET (2014)
6. Sonwane, A.D., Bhonge, V.N., Khandare, A.: Design and development of wireless sensor node for anti-poaching. IEEE Xplore (2016)

Chapter 17
Performance Issues and Monitoring Mechanisms for Serverless IoT Applications—An Exploratory Study



Shajulin Benedict

Abstract Performance monitoring and analysis of applications are imperative tasks for application developers, including serverless application developers. Application developers, in general, focus on developing their application logic rather than looking into the performance bottlenecks of serverless functions. In this review article, available performance analysis tools and monitoring mechanisms for IoT-enabled serverless applications are explored. Efforts are taken to investigate the existing performance analysis tools, their merits, demerits, and the performance metrics for IoT-enabled serverless applications. Further, various challenges, highlighting the need for the attributes for implementing sophisticated performance analysis tools of IoT-enabled serverless applications, are discussed. This article will be beneficial for the performance analysis tool developers or the serverless cloud application developers.

17.1 Introduction

IoT-enabled application development has gained a wide attraction in various domains such as transportation, agriculture, industry, smart cities, health care, smart buildings, and so forth. The application of serverless functions, empowered by IoT, cloud, and edge computing, has emerged as a profiting computing paradigm in the recent past. Developers have utilized serverless functions to enable an event-driven IoT programming aspect; to niche functions on the pathway of service flows; to trigger stream processing elements; and to bridge an integrated edge-cloud learning framework.

Although serverless computing had exhorted the minds of cloud and IoT application developers, it appears to have been subdued in the utilization expectations owing to the performance concerns and poor practical knowledge of the users. A few

This work is partially funded by IIITKottayam Faculty Research fund and AIC-IIITKottayam project.

S. Benedict (✉)
Indian Institute of Information Technology Kottayam, Kottayam, Kerala 686635, India
e-mail: shajulin@iiitkottayam.ac.in
URL: https://www.sbenedictglobal.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_17

notable performance concerns of serverless applications include (i) resource mapping issues, (ii) non-availability of resources, (iii) slow start and warm start issues, (iv) memory allocation issues, (v) security concerns, and so forth. Obviously, tools need to be assisted for pinpointing the performance bottlenecks of these applications.

This paper reveals the multi-cornered performance issues that underscore IoT-enabled serverless applications; explores a generic IoT-enabled serverless architecture for implementing serverless applications; and, lists the merits and demerits of the available performance analysis tools for serverless-oriented IoT-enabled applications. Besides, the performance metrics that are specific to serverless applications and the future research directions are highlighted in the article.

The rest of the paper is discussed as follows: Sect. 17.2 highlights the related works, Sect. 17.3 introduces the generic architecture of IoT-enabled serverless computing environment; Section 17.4 delves into the key performance concerns of IoT-enabled serverless applications; Section 17.5 discloses the tools and their challenges; and, Sect. 17.6 provides a few conclusions and outlooks.

17.2 Related Works

Research works on developing IoT cloud applications have much evidenced the utilization of serverless functions in the recent past owing to the resource efficiency or scalability features. For instances, Gusev et al. [1] have proposed the application of serverless functions on IoT devices to increase the scalability of streaming applications; Razin et al. [2] have applied serverless computing for green oil gas industry. Apart from IoT cloud-specific applications, serverless functions have also become a pristine concept for blockchains [3] and programmers [4–6].

There exist a few platforms for implementing serverless functions and hosting them. Such platforms are available in two dimensions: Opensource and commercial facets–platforms such as Kubeless, OpenWhisk, Fission, OpenFaaS, and so forth are opensource; and, Platforms such as AWS Lambda, Google Cloud Functions, Microsoft Azure Functions, and so forth are commercial platforms for executing serverless functions. Mohanty et al. [7] have analyzed the efficiency of such opensource serverless platforms in terms of the scalability and the response time of handling multiple client requests; similarly, Lee et al. [8] have revealed the efficiency of different serverless platforms while executing the dynamic workload functions on them.

In general, performance monitoring tools could pinpoint the performance problems of applications [9, 10]. Although there are a few performance analysis tools that discuss on the performance concerns of serverless functions, a comprehensive study that exposes the intricacies of those performance monitoring tools and those tools which highlight the future research directions are very rarely available in the literature. Standalone discussions on analyzing the impact of executing serverless functions on the available computing platforms exist. For instance, Eric et al. [11] have manifested the utilization of suitable communication patterns for serverless

environments. This paper disclosed the inner details of the available performance tools of serverless functions when executed on serverless platforms.

17.3 Serverless IoT-Working Model

Serverless computing is an event-driven-based execution model that advances in cost ownerships; i.e., the functions are executed on servers only when they need to be executed. Long-serving server applications could benefit from this computing model as the server functions seek available resources only during the executions. Such functions are often individually managed functions. Most commonly utilized serverless functions are single-task functions such as accessing a database, registering user credentials into the database, functioning actuators, invoking a periodic task, scheduling a task, and so forth.

A pictorial representation of the IoT-enabled serverless architecture, where serverless functions could be augmented at four levels of IoT-enabled applications, is shown in Fig. 17.1. The modes of operations or the possible functionalities of serverless functions at each level of the compute nodes are discussed below:

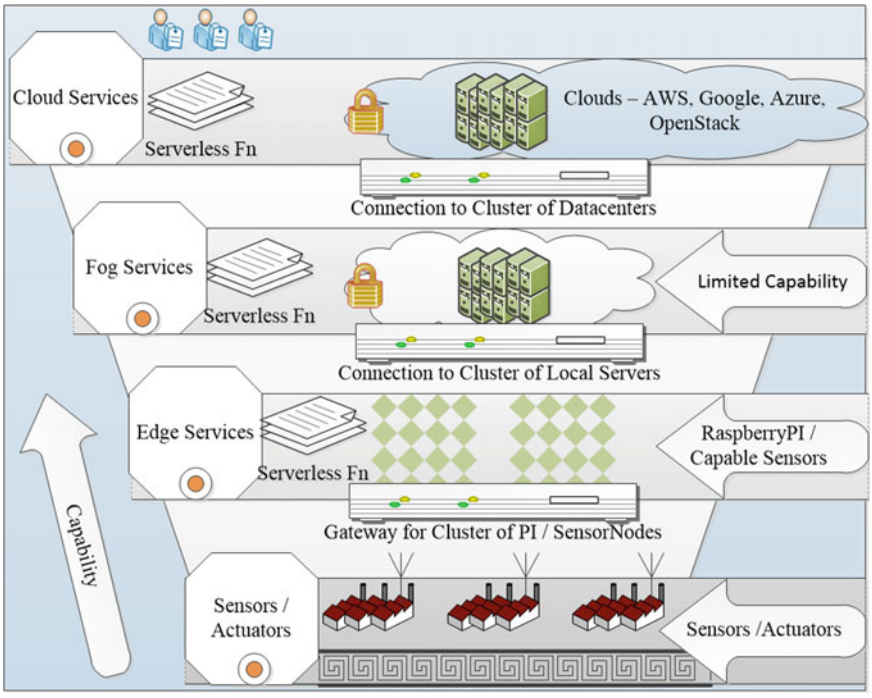


Fig. 17.1 Generic architecture of IoT-enabled applications—the suitability of serverless functions

1. *Cloud level*: Serverless functions are more predominately applied at the cloud level computing machines for various purposes: (a) invoking security flavors, (b) executing data analytic algorithms, (c) performing performance monitoring of cloud resources and applications, (d) dealing with user interfaces, (e) enriching the business capabilities of applications, (f) enabling logging services, and so forth.
2. *Fog Level*: Understanding the performance concerns of transferring all sensor data to the cloud environments and processing them at the cloud, a few compute-capable devices such as servers or VMs are utilized to assist the IoT applications at the fog level. The responsibilities of these fog nodes are (i) to contextually filter sensor data (if required), (ii) to perform inter-operable services between heterogeneous sensor networks, (iii) to analyze the sensor data based on algorithms, including Machine Learning (ML) algorithms, and so forth.
3. *Edge Level*: Typically, IoT sensors or actuators are connected to edge nodes. The edge nodes are power constraint nodes such as raspberry pi nodes or IoT gateways or pi clusters. The edge-level serverless functions are utilized to trigger sensor nodes or actuator nodes based on the tasks levied in the functions; and, to deal with the fault-tolerance issues of sub-tasks of IoT applications.
4. *Sensor Level*: It is an arduous task to launch serverless functions at the sensor level of IoT applications. IoT sensors or actuators are placed at this level. These devices are, mostly, power-constrained or battery-operated sensor nodes that are capable of collecting sensor data (measurable property such as temperature, humidity, or so forth); and, actuators that are capable of performing actions based on the control signals.

A general practice of executing serverless functions on all these levels is to utilize containers/orchestration platforms on servers. For instances, kubeless functions, iron functions, Fission, Fn platform, and so forth are executed on Kubernetes; OpenWhisk functions are executed on IBM Apache servers; AWS Lambda and Sparta are executed on AWS instances; Google Cloud Functions are accomplished on Google Compute Engine instances; and, Microsoft Azure functions are executed on Azure. To process applications nearer to the data sources and to preserve the data privacy, a few IoT-enabled platforms hierarchically offer their services. For instance, Amazon has extended the AWS IoT product with specific capabilities in the AWS GreenGrass product. The AWS GreenGrass product enables cloud services on edge devices. Services such as AWS Lambda, Amazon Simple Storage Service (S3), Amazon Kinesis, Amazon Simple Queue Service (SQS), and so forth, are executed on edge nodes.

17.4 Performance Monitoring Mechanisms

This section describes the performance monitoring mechanisms and the existing performance issues or challenges of involving serverless functions on IoT-enabled cloud-based applications.

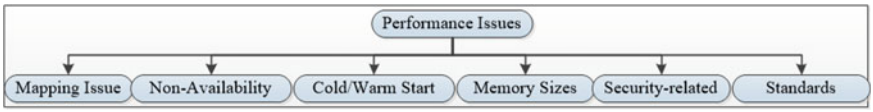


Fig. 17.2 Performance concerns of serverless functions

Performance Issues—Serverless Functions A list of performance issues that influence the efficient working of serverless functions on cloud/edge-level resources is discussed below (see Fig. 17.2).

Mapping issue: Often, the serverless functions are not exactly mapped to the available heterogeneous IoT sensors or edge/fog/cloud resources. The protocols and rules of different sensors belonging to various companies should be compatible so that the sensors are connected each other [4, 12].

Non-Availability: The execution-related resources such as containers, memory, and network should be readily available for keeping serverless functions active in real-time irrespective of their mobility. However, the solidity of availability is beleaguered owing to the provider issues or the issues relating to creating VMs. To improve the availability of functions, retrospection of previous states is provided in some architectures.

Cold or Warm Start Issues: The first time instantiation of VMs or containers for executing serverless functions, also named as `Cold Start` of serverless functions, will certainly draw down the execution time of serverless functions when they are executed on edge environments. Similarly, if the containers are not bootstrapping for the first time, then the image is assumed to be powered ON which is commonly termed as `Warm Start`. In fact, the `Cold Start` leads to a high overhead and the `Warm Start` arrives at high energy consumptions [13, 14].

Impact of Memory Sizes and BigData: Most of the serverless functions offered by providers, such as AWS, Google, or Microsoft, have options to specify the memory sizes—for instance, AWS lamba offers one million requests for function invocations at the free of cost with memory sizes ranging from 128 to 3008 MB. Increasing the memory sizes would extensively increase the CPU instantiation time of applications [15].

Security Concerns: IoT applications, in general, are prone to security issues such as device hacking, authentication issues, cross site scripting, injection attacks, file manipulation, and so forth.

Standards, Protocols, Programming Languages: Lack of proper standards or protocols for serverless functions at the IoT domain has also degraded the performance of IoT-enabled serverless applications. Besides, the utilization of apt programming languages for serverless functions is also an important parameter to improve the performance of IoT applications—for instance, services or functions written in golang or nodejs suffers from a hefty start time due to their inherent dependencies.

Monitoring Mechanisms The possible ways of monitoring serverless functions include: (i) *Tracing Approach:* Tracing based monitoring tracks serverless functions when executed on edge/cloud machines. It is carried out by sniffing into the log files or error messages of serverless functions in a distributed fashion—often, named as *distributed tracing*; (ii) *Sandboxing Approach*: In the sandboxing approach, an isolated and secure environment will be created to study the performance impact of the newer services or functions on computing resources. IoT applications, especially automated vehicles or industry-oriented IIoT applications, require continuous operations or predictive maintenance/allotment [16, 17] to fulfill the automated functioning of machines to enrich the supply-value chain of the upcoming innovative products. Adopting a newer service or a beta-version of data analytic algorithms might lead to a disfunction of the whole machinery. In order to avoid the performance impact of adding such services directly at the production time, the sandboxing approach of performance analysis is recommended; (iii) *Modeling/Prediction Approach*: To avoid performance monitoring overheads, modeling of the end-to-end execution of serverless functions could be a suitable approach for getting the performance insights of functions. For instance, authors of [18] have studied the performance of serverless applications by modeling the dependencies between serverless functions and their associated memories, storage, other computing units; and, (iv) *Language-Inherent or Instrumentation Approach*: The language-inherent or instrumentation approach of monitoring the performance of serverless functions will have the capability of injecting a few predefined functions or notations into the serverless functions before their executions.

17.5 Monitoring Tools—A Comparative Study

The nine most commonly utilized performance analysis tools for monitoring serverless functions in IoT-enabled edge/fog/cloud computing levels are Amazon X-Ray (T1), Google StackDriver (T2), ApplicationInsights (T3), CloudWatch (T4), Prometheus (T5), Thundra (T6), AppDynamics (T7), Wavefront (T8), and Epsagon (T9). A synopsis of these tools are discussed below:

1. `T1 -- Amazon X-ray` applies the instrumentation features to monitor Lambda functions or services. It is a distributed tracing tool to analyze the performance issues of serverless functions, such as, average latency, utilization chart for databases, function completion ratio (Failure rate), and so forth, when executed on the Amazon Lambda platform;
2. `T2 -- Google-Stackdriver` offers several preferences such as *Stackdriver Debugger*, *Stackdriver Trace*, and *Stackdriver Profiler*, in the form of the performance monitoring API to the Google Cloud users;
3. `T3 -- ApplicationInsights` service is most commonly utilized for analyzing the serverless functions of Microsoft Azure namely *Azure Functions*. It requires minimal configurations for delving into the performance details of functions;

4. T4 -- CloudWatch is event-based monitoring tool for AWS Lambda functions;
5. T5 -- Prometheus, combined with the Grafana tool, is a combinatory open-source tool utilized for measuring the performance issues of serverless functions namely (i) Kubeless functions on Kubernetes or (ii) Fx functions of Apache OpenWhisk on container engines;
6. T6 -- Thundra tool provides an asynchronous mode of monitoring of the serverless functions with a provision for including the automated instrumentation feature. It offers an end-to-end monitoring support for serverless applications when executed on heterogeneous computing nodes. The tool is claimed to provide a continuous configuration, enforcement, and verification of security and controls;
7. T7 -- AppDynamics is an application performance management system that analyzes the entire application performance of serverless-enabled applications;
8. T8 -- WaveFront and Delta-WaveFront tools are the products of VMWare that monitor the performance of AWS Lambda serverless functions with a special feature to sum up all metrics together;
9. T9 -- Epsagon tool, an offering from a company with different pricing models, focusses on vividly highlighting the issues relating to the data pipelines of Lambda functions emerging from various IoT sensors. Most preferably, this tool deals with the devices belonging to one organizational domain or docker images.

A comparison of these nine performance analysis tools of serverless functions is discussed in different perspectives: (i) commercial versus open source, (ii) monitoring approach, (iii) expressiveness, (iv) measurement granularity, and (v) solution availability (See Table 17.1). In addition, the capabilities of these tools are given in Table 17.2.

17.5.1 Challenges of Tools—Research Directions

Although performance analysis tools are available in the market, there are several pitfalls which could be further improved in various directions as discussed below: (i) The serverless functions are often short-lived. Naturally, it becomes an arduous task for any performance analysis tools to reproduce the same fault or to trace the same behavior of serverless functions; (ii) Delving into the large volume of performance data available in the log files is a challenging task for the existing tools. Although a few tools visit the log files, they are unable to find the root cause of the performance problem within a short span of time or to depict the faults using sophisticated *visualization* options; (iii) In line with the problems faced by serverless functions of IoT applications, the associated performance analysis tools still experience the issue of vendor lock-in issues. This is a serious concern while considering the realization of a federated IoT-enabled cloud application where the performance analysis tool has to deal with the pros and cons of multiple cloud providers; (iv) Most

Table 17.1 Comparison of performance analysis tools for serverless applications

Description	T1	T2	T3	T4	T5	T6	T7	T8	T9
Commercial	✓	✓	✓	✓		✓	✓	✓	✓
Opensource					✓				
FreeTrial Version	✓	✓	✓	✓	✓		✓	✓	✓
Logs		✓		✓			✓		
Plugin			✓						
Tracing	✓				✓	✓	✓	✓	✓
Library		✓			✓	✓			✓
Daemon/Agent			✓				✓		
Service	✓			✓				✓	✓
Function	✓	✓	✓		✓	✓		✓	
Task		✓	✓						
Application				✓		✓	✓		✓

Table 17.2 Performance monitoring tools—important features

Tools	Measuring devices	Features
T1	Amazon LambdaFunctions, EC2	Allows data annotation
T2	GCE, gateways	Profile approach, distributed tracing
T3	Microsoft Azure Functions	Custom metrics
T4	EC2	Specific to hosts
T5	Containers, nodes, and clusters	Kubernetes cluster based support
T6	Amazon LambdaFunctions, EC2	Asynchronous tracing, low overhead
T7	Nodes, containers	Agent based monitoring
T8	Amazon LambdaFunctions, EC2	Less overhead
T9	Amazon LambdaFunctions, EC2	One organization focus on data pipelines

of the available tools endeavor the integration issues of different sensors or nodes belonging to various vendors. It is a mandatory task for the tools to ensure that they are capable enough to survive such environments while measuring the performance data of the functions; and (v) The existing tools are not capable to precisely identify the energy consumption of serverless functions when executed on federated cloud environments.

17.6 Conclusion

The emergence of IoT-enabled applications has increased in the recent past in various sectors such as smart transportation, smart cities, industry, health care, and so forth. The application of serverless functions for executing such applications improved the performance efficiency of them in terms of metrics such as scalability, latency, and the number of requests handled by services. Although practitioners have continually urged for the performance monitoring tools that could assist them to pinpoint the performance bottlenecks of serverless functions, there exists a very few tools that illustrate the relevant performance metrics. This paper explored the emerging performance concerns of serverless functions that are specific to IoT-enabled serverless applications. In addition, it has given deeper insights into the tools for monitoring serverless functions. The challenges and research directions were also provided for the researchers to focus on in the near future while developing IoT-enabled serverless-based applications or the associated performance monitoring tools.

References

1. Gusev, M., Koteska, B., Kostoska, M., Jakimovski, B., Dustdar, S., Scekic, O., Fahringer, T.: A deviceless edge computing approach for streaming IoT applications. IEEE Internet Comput. **23**(1), 37–45 (2019)
2. Hussain, R.F., Salehi, M.A., Semiari, O.: Serverless Edge Computing for Green Oil and Gas Industry (2019). arXiv:1905.04460v1
3. Chen, H., Zhang, L.: FBaaS: Functional Blockchain as a Service. In: Lecture Notes in Computer Science, pp. 243–250 (2018). https://doi.org/10.1007/978-3-319-94478-4_17
4. Bhattacharjee, A., Barve, Y., Khare, S., Bao, S., Gokhale, A.: Stratum: A Serverless Framework for the Lifecycle Management of Machine Learning-based Data Analytics Tasks. arXiv:1904.01727v1 (2019)
5. Hung, L., Kumanov, D., Niu, X., Lloyd, W., Yeung, K.Y.: Rapid RNA sequencing data analysis using serverless computing (2019). https://doi.org/10.1101/576199
6. Werner, S., Kuhlenkamp, J., Klems, M., Muller, J., Tai, S.: Serverless big data processing using matrix multiplication as example. In: 2018 IEEE International Conference on Big Data (Big Data) (2018). https://doi.org/10.1109/bigdata.2018.8622362
7. Mohanty, S.K., Premsankar, G., Di Francesco, M.: An evaluation of open source serverless computing frameworks. IEEE CloudCom **2018**, (2018). https://doi.org/10.1109/cloudcom2018.2018.00033
8. Lee, H., Satyam, K., Fox, G.: Evaluation of production serverless computing environments. In: IEEE 11th CLOUD'18 (2018). https://doi.org/10.1109/cloud.2018.00062
9. Benedict, S., Rejitha, R.S., Bright, C.: Energy consumption analysis of HPC applications using NoSQL database feature of energyanalyzer. In: ICC 2014, vol. 8993. Springer, LNCS (2014). https://doi.org/10.1007/978-3-319-19848-4_7
10. Benedict, S., Gerndt, M.: Automatic performance analysis of OpenMP codes on a scalable shared memory system using periscope. In: PARA 2010, vol 7134. Springer LNCS (2010). https://doi.org/10.1007/978-3-642-28145-7_44
11. Eric et al. (2019) https://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-3.html. Accessed in June 2019
12. Aumala, G., Boza, E.F., Ortiz-Aviles, L., Totoy, G., Abad, C.L.: Beyond load balancing: package-aware scheduling for serverless platforms. In: 19th IEEE/ACM International Symposium on CCGRID, pp. 282–291 (2019)
13. Lin, P.-M., Glikson, A.: Mitigating Cold Starts in Serverless Platforms A Pool-Based Approach (2019). arXiv:1903.12221v1
14. Wan, J., Han, S., Zhang, J., Zhu, B., Zhou, L.: An image management system implemented on open-source cloud platform. IEEE IPDPS **2013**, (2013). https://doi.org/10.1109/ipdpsw.2013.176
15. Klimovic, A., Wang, Y., Kozyrakis, C., Stuedi, P., Pfefferle, J.: A Trivedi understanding ephemeral storage for serverless analytics. In: Proceedings of the 2018 USENIX Conference, pp. 789–794 (2018)
16. March, S.T., Scudder, G.D.: Predictive maintenance: strategic use of IT in manufacturing organizations. Inform. Syst. Front. **21**(2), 327–341 (2017). https://doi.org/10.1007/s10796-017-9749-z
17. Sezer, E., Romero, D., Guedea, F., Macchi, M., Emmanouilidis, C.: An Industry 4.0-enabled low cost predictive maintenance approach for SMEs. In: 2018 IEEE ICE/ITMC (2018). https://doi.org/10.1109/ice.2018.8436307
18. Winzinger, S., Wirtz, G.: Model-based analysis of serverless applications. In: 2019 IEEE/ACM 11th International Workshop on Modelling in Software Engineering (MiSE), pp. 82–88 (2019). https://doi.org/10.1109/MiSE.2019.00020

Chapter 18
Detection of Pemphigus Vulgaris in Development Stage of Skin Erosion



Mayuri Vagh and Dipali Kasat

Abstract Pemphigus vulgaris (PV) is an autoimmune blister disease which is a common type of pemphigus. It affects the oral mucosa and skin area. PV blister looks the same as the other blister disease. Thus, diagnosis of PV disease is difficult because of the fact that only one characteristic is identifying the PV blister that is Nikolsky's sign (bulla spread sign). Ruptures of the blisters make erosion on the surface of the skin area. The clinical diagnosis of PV is more painful. Therefore, selecting an area of relevant diagnosis through image processing is important not just for dermatologists but for non-expert person as well. The proposed work is an attempt to provide a solution for the detection of PV in the developmental stages of skin erosion. We developed a systematic and effective framework for identifying PV from relevant information, image data, and the description. Our system works in two sections, the first is interrogation about symptoms and the second to detect blister images. The accuracy rate of the system to successfully detect PV blister is 88.67%.

18.1 Introduction

Blister (Bulla) skin diseases are a group of autoimmune disorders. Pemphigus vulgaris (PV) is one of the blister diseases. The life-threatening PV diseases affect 1–5 patients per 1,000,000 populations per year [1]. PV can be proven to be a serious fatal disease if it is untreated. The majority of occurrences are found between 40 and 60 years of age [2]. The occurrence ratio in women is higher compared to men and the ratio range from 1:2 [3]. In 70–90% of the cases, it presents itself with painful and long-persisting erosion and blisters of the mucous membranes, especially of the oral mucosa that spread to other parts of the body area in its developmental stages [4]. It has an important symptom which is bulla spread sign. In the active phase of PV, it is characterized by the epidermal detachment caused by the pressure at the edge

M. Vagh · D. Kasat (✉)
Computer Engineering Department, Sarvajanik College of Engineering and Technology, Surat, Gujarat, India
e-mail: dipali.kasat@scet.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_18



Fig. 18.1 PV blister image
Source by Dermatologist

of a blister, and it is also known as Asboe—Hansen sign or Nikolsky's sign [5]. The proposed system provides a non-probing image-based approach for the detection of PV blisters (Fig. 18.1).

18.2 Literature Survey

Schlesinger et al. in their study, "Nail involvement in pemphigus Vulgaris [6]," on 64 patients with PV disease, found that nail changes in 30 out of 64 PV patients, with 14 patients having biopsy-proven PV of the nail. Habibi et al. discussed "Nail changes in pemphigus vulgaris [7]," found that 25 out of 79 PV patients had nail changes during the occurrence of PV disease.

In 2017, Ansari and Sarode, in their work "Skin cancer detection using image processing [8]" used features index information of data present in an image. In 2016, Kumar et al. in their work "Dermatological disease detection using image processing and machine learning [9]" did mapping with the ROIs for skin analysis. It is compared with different classification techniques.

In 2015, Amarathunga, et al. showed "Expert system for diagnosis of skin disease [10]" and performed diagnosis of skin disease by the expert system. The author used the fuzzy logic controller to generate the result with symptoms. In 2018, Zomnrylü, et al., present "The quantitative assessment of progressive dermatologic manifestations in selected ROI (Region of Interest) [11]" developed the algorithm for detecting the blister lesions. It represents a unique attempt to develop a technique of detecting targeted blister lesions in affected skin areas. From the literature survey, we have identified some challenges:

The color after the burst of PV blister is identical to the burnt color and visually the same as the PV blister and the erosion. PV blisters are identical to other autoimmune skin blisters [7]. All autoimmune blisters are visually the same as another blister and covered by the fluid-filled bubble. The identification of significant symptoms of a unique key factor is called Nikolsky's sign [5].

18.3 System Overview

Figure 18.2 depicts a system to decide PV or Not-PV disease based on a flow diagram. The proposed system is divided into two parts: The primary section is an expert system, and the second section is image analysis. In the primary section, we designed the questions for interrogation of symptoms based on some criteria. The next question is determined based on the answers given by the patient to the required questions. In the secondary section, the image analysis consists of preprocessing, segmentation, extract ROI, and feature extraction followed by classification of the blister as "PV" or "Not-PV."

18.3.1 Primary Section/ Interrogate About Symptoms

The primary section is "Interrogate About symptoms" with the expert system. All questions are based on PV symptoms. The questions depend on the answer to the previous question. The system performs diagnosis for PV patients by asking questions related to the conditions that require objective answers.

Figure 18.3 depicts the weightage assigned to answers of the questions asked based on study of symptoms observance in 100 patients. If symptoms are high chance in PV, then the option is more prevalent and carries more weightage. If symptoms show less chance of PV, then the option is less prevalent and carries a lower weightage. The system calculates the weightage at the end of Q/sessions based on the weightage

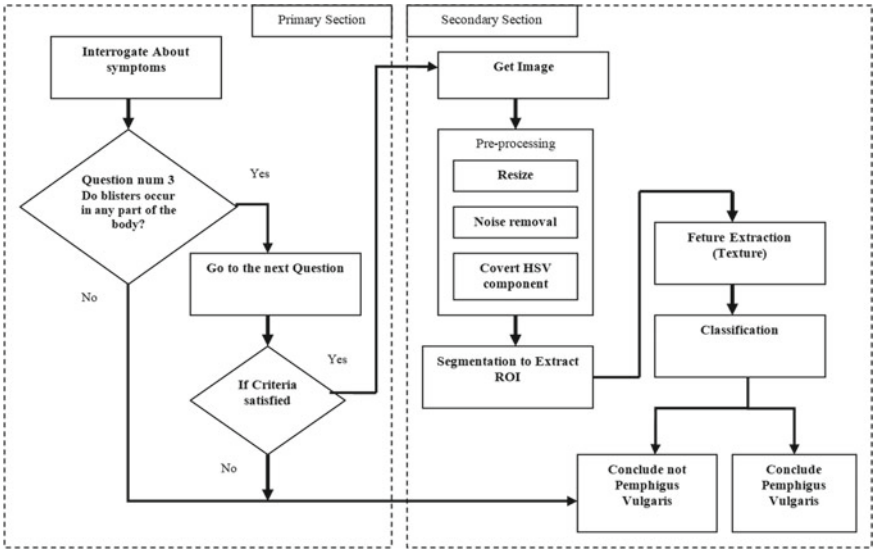


Fig. 18.2 Flow diagram

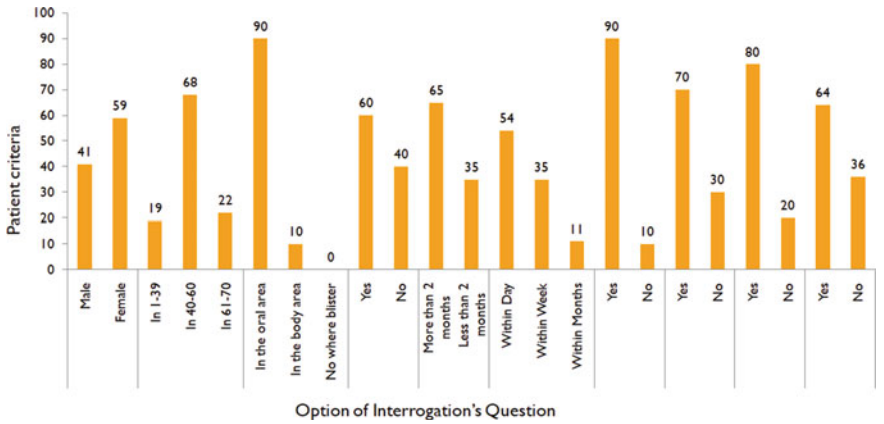


Fig. 18.3 Representation of patient ratio

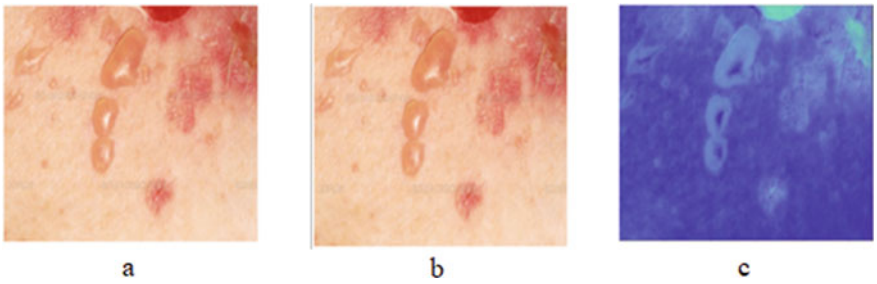


Fig. 18.4 **a** Original image, **b** use resize and noise removal operation, and **c** convert into components of HSV components that are hue, saturation, and intensity

of each answer and deducts the symptoms match result. If the system match is not satisfied with PV disease criteria, then there is no point in going further into the image processing section, but if the criteria are satisfied, then the image analysis part is executed (Fig. 18.4; Table 18.1).

18.3.2 Secondary Section/ Image Analysis

After the criteria get satisfied, the next step is the image analysis of the blister image given by the patient. The patient can submit the image in two ways, capture images using a camera and from the database. After the "get image" process, preprocessing is the primary step. Every image is not in the precise size so, we resize into the standardized size of 400*400. We use Gaussian blur filtering for noise removal in image preprocessing. For the color spaces differentiate the image between hue,

Table 18.1 Questionnaire of interrogation about symptoms

No.	Question	Options	Weight
1	The gender of the patient	Male	8
		Female	10
2	What is the age of the patient?	Age 1–39	8
		Age 40–60	10
		Age 61–70	2
3	Where blisters first appeared?	In the oral area	10
		In the body area	2
		No blister	0
4	Whether blisters are present in the stomatitis area	Yes	10
		No	8
5	A period of blisters in the oral area	More than 2 months	10
		Less than 2 months	2
6	How much time taken for rupturing of blisters	Within day	10
		Within week	8
		Within months	2
7	Whether flaccid arose on healthy skin while rupturing blisters?	Yes	10
		No	2
8	Having itching problem or not?	Yes	2
		No	10
9	Nails are affected?	Yes	10
		No	8
10	Is there any maceration in skin folds?	Yes	10
		No	8

saturation, and intensity components. The saturation component gives the best result. After the saturation color component uses the segmentation method. It is simple and effectively isolates objects in the image. With the help of a threshold, the blister will appear in the foreground. Then, we apply the contour. It will join the line to all the points along the boundary of the image that are having the same intensity. The contour draws the boundary line of the blister shape. We set the region of interest (ROI); it does samples identified for a selected purpose. That blister will be detected with the ROI. Figure 18.5 shows the final results generated by the application with the extracted blister highlighted in green.

The next step is extracting the useful feature which can classify the input image as PV or Not-PV. Important features are color, shape, and texture for the detection of PV blister. We have extracted texture features from the image for classifying the input image as PV or Not-PV. The next step is to classify it as PV or Not-PV. We propose an improvement of classification with the random forest method.

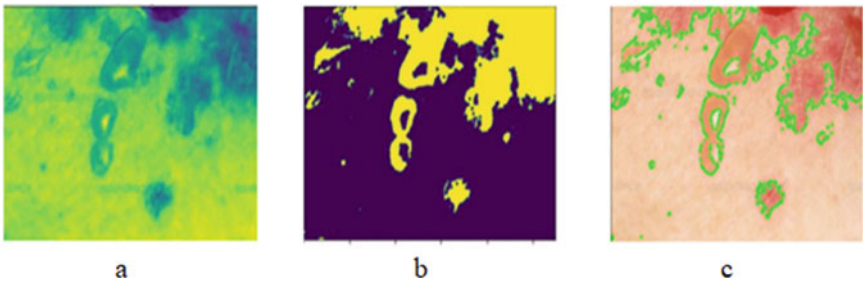


Fig. 18.5 **a** Select saturation component, **b** segmentation, and **c** extracted blisters (ROI)

18.4 Result Analysis

We collected images from different websites of specific blister diseases and the medical expert. The database has 330 images (164 PV blister images, 185 Non-PV). We worked in two different sections: The first section is interrogation about symptoms as the primary section, and the second section is image analysis. In the primary section, the patient will answer the questions from experiences about the blister period. If the criteria of the patient's selected options are calculated to be 50% or less than 50%, then the patient does not have symptoms of PV, so we will suggest a biopsy. And if the criteria of the patient's selected options are calculated to be more than 50%, then the patient has symptoms of PV, so we will suggest uploading the image of blister for the image analysis. After the primary section, the second section is the image analysis. These are the steps: get image (input image), preprocessing, segmentation, feature extraction blister analysis, and the classification of PV blister. For the right accuracy and types of errors, we calculate a confusion matrix. It contains information on the actual and predicted value of the classifier. We get 133 images correctly classify with RF classifier out of 150 images (Table 18.2).

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positives} + \text{False Positive} + \text{True Negative} + \text{False Negative}} \tag{18.1}$$

We classify the image as PV or Not-PV using the RF, SVM, and DT. SVM model sets the clear margin, suitable for a small data set and separation between classes. DT model is a versatile, specific value to each problem, easy to use. Figure 18.6

Table 18.2 Confusion matrix RF, SVM, and DT

	Actual					
	RF		SVM		DT	
*P*rediction	64	10	45	29	62	12
	7	69	14	62	11	65



Fig. 18.6 Overall classifier's performance

embarks that RF classifier will handle the missing values and maintain the accuracy of a large amount of data. It would not allow overfitting in the model. RF has a maximum overall accuracy of 88.67% and proved to be better among all for our work. Our system has been tested with different scenarios. The first scenario is if the system does not get any true skin color in the image, then the system not going into the further image processing step. For example, a user uploads a random image without any skin color. The second scenario is if criteria are satisfied but with the user uploading the wrong image (the image blisters are absent). For example, the image is human skin. But the image has no blister. For the PV detection, blister is compulsory. The system will start the further process, but the system will identify the image at the ROI step. The blisters are not available in the image (Table 18.3).

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} \tag{18.2}$$

Table 18.3 Recall and precision of incorrect input; recall and precision of submit various quality of image

Scenario	Total number of Image	Recall	Precision	Accuracy
Recall and precision of incorrect input				
Not submit the human skin image	20	0.88	0.88	0.9
Submit wrong image after satisfying the criteria	25	0.9	0.83	0.88
Recall and precision of submit various quality of image				
Submit various types of noise image	21	0.94	0.1	0.95
variation in image resolution	25	0.9	0.83	0.88

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} \tag{18.3}$$

In the third scenario, the images have various types of noise. By noise in image pixel values that do not reflect the true intensities of the real pixels. But our system classifies the correct image. In the fourth scenario, the patient gives a variation in image resolution. Image pixels seem to blend together to form the image. The system gets glitch images and loses the important pixels that still classify the correct image.

18.5 Conclusion

PV is one of the rare autoimmune blister diseases. PV blisters are painful and develop on the skin, mouth, nose, throat, and genitals. In this paper, we propose a system for the detection of PV. The system works in two sections, the primary section interrogates about symptoms, and the secondary section performs an analysis of the image to classify it as a PV or non-PV image. In classification, phase one is training and the second is the testing phase. Out of 330 total images, 180 images are considered in the training phase, and the remaining 150 images are considered for the testing phase. The resulting analysis proves the 88.67% accuracy obtained. The proposed system can assist non-technical and technical persons to detect the PV blister. The system can be extended for the detection of the other blister disease in the future.

References

1. Shamim, T., Varghese, V.I., Shameena, P.M., Sudha, S.: Pemphigus vulgaris in oral cavity: Clinical analysis of 71 cases. Med. Oral Pathol. Oral Cir. Buccal. **13**(10), 622–666 (2008)
2. Iamaroon, A., Boonyawong, P., Klanrit, P., Prasongtunskul, S., Thongprasom, K.: Characterization of oral pemphigus vulgaris in Thai patients. J. Oral Sci. **48**(1), 43–46 (2006)
3. Tamgadge, S., Tamgadge, A., Bhatt, D.M., Bhalerao, S., Pereira, T.: Pemphigus vulgaris. Contemp. Clin. Dent. **2**(2), 134–137 (2011)
4. Baum, S., Sakka, N., Artsi, O., Trau, H., Barzilai, A.: Diagnosis and classification of autoimmune blistering diseases. Autoimmun. Rev. **13**, 482–489 (2014)
5. Porro, A.M., Seque, C.A., Ferreira, M.C.C., Enokihara, M.M.S.E.S.: Pemphigus vulgaris. Anais brasileiros de dermatologia (2019)
6. Schlesinger, N., Katz, M., Ingber, A.: Nail involvement in pemphigus vulgaris. Br. J. Dermatol. **146**(5), 836–839 (2002)
7. Habibi, M., Mortazavi, H., Shadianloo, S., Balighi, K., Ghodsi, S.Z., Daneshpazhooh, M.: Nail changes in pemphigus vulgaris. Int. J. Dermatol. **47**, 1141–1144 (2008)
8. Bano Ansari, U., Sarode, T.: Skin cancer detection using image processing. Int. Res. J. Eng. Technol. **04** (2017)
9. Kumar, V.B., Kumar, S.S., Saboo, V.: Dermatological disease detection using image processing and machine learning. In: Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), Lodz, Poland, 19–21 (2016)

10. Amarathunga, A.A.L.C., Ellawala, E.P.W.C., Abeysekara, G.N., Amalraj, C.R.J.: Expert system for diagnosis of skin diseases. Int. J. Sci. Technol. Res **4**(1), 174–178 (2015)
11. Zeljkovic, V., Druzgalski, C., Mayorga, P., Tameze, C.: Quantitative assessment of progressive dermatologie manifestations in selected ROI (Region of Interest), Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE), Mexico, 20–25 (2018)

Chapter 19
Anomaly-Based Intrusion Detection System in Two Benchmark Datasets Using Various Learning Algorithms



Thongam Jayaluxmi Devi and Khundrakpam Johnson Singh

Abstract The research in network intrusion detection has escalated since past few years. There are various methods and systems being proposed related to intrusion detection. But the changing nature of attack leads to the need to deal with every possible aspect to increase the detection efficiency. The anomaly-based intrusion detection has always been in the limelight because of its detection capability of the unknown pattern. In the study, feature reduction is performed using gain ratio, correlation, information gain and symmetrical uncertainty, and the selected features are used to train the machine learning techniques such as Naive Bayes classifier, sequential minimal optimization (SMO), J48 classifier and random forest. The KDD cup 99 and NSL-KDD datasets were used as a benchmark reference in building the models. The experiment was carried out with an aim to compare the performance of various classifiers between the two datasets. Results show that the feature reduction improves the detection efficiency and can obtain accuracy as high as 99.91%.

19.1 Introduction

The evolution in Internet technologies with concurrent advancement in the number of network attacks has led to the escalation of the importance of research in network intrusion detection. Despite the phenomenal improvements in the area of research, still there are various ways to improve and enhance detection and prevent the network-based attacks. According to [1], "threat is a potential possibility of a deliberate unauthorized attempt to (i) access information, (ii) manipulate information or (iii) render a system unreliable or unusable." The present world solely depends on technology. And with time almost all data are stored in the cloud, and every work is performed online. As such, it becomes a necessity to secure our data or system from intrusions. "The term anomaly-based intrusion detection in networks refers to the problem of finding exceptional patterns in network traffic that do not conform to the expected normal behavior"[2]. "The fundamental concept behind this technique is to define

T. J. Devi (✉) · K. J. Singh
Department of CSE, NIT Manipur, Imphal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_19

the behavior of the network and/or system, where this predefined behavior is then compared with the normal behavior. The result will be either to accept it or to trigger alarm management system for further investigation"[3].

According to [4], we have two main classifications of intrusion detection system: signature-based and anomaly-based. Signature-based includes exploring network traffic for any malicious bytes or packet sequences. It is easy to establish and acknowledge if we know the behavior of the incoming network. The problem arises when we do not have the pattern of the incoming traffic. In order to increase its validity, we need to update and upgrade our model for any new attack pattern. The anomaly-based intrusion detection system need not know the pattern of the attack in order to alert the system. By defining a normal behavior, the anomaly-based IDS is ready to operate. Any observation of deviation from the normal behavior or the pattern will give an alarm to the system for intrusions.

The rest of the paper consists of related works, datasets, proposed methodology, results and discussion and conclusion.

19.2 Related Works

Table 19.1 shows some of the advantages and limitations of various surveys.

19.3 Datasets

Two datasets have been used in the study: KDD cup 99 and NSL-KDD.

19.3.1 KDD Cup 99

The KDD cup 99 dataset was developed at the MIT Lincoln Laboratories. It is classified into four main attacks: denial of service (DOS), remote to local (R2L), user to root (U2R) and probing. KDD dataset consists of 41 features and one label which specifies the attack name. Table 19.2 shows the features of KDD cup 99.

19.3.2 NSL-KDD

In the study of KDD, two major problems arise, i.e., redundant and duplicate records in training and testing datasets, respectively, which tremendously affect the performance of evaluated systems that leads to the introduction of NSL-KDD.

Table 19.1 List of surveys

Authors	Datasets	Methods	Advantages	Limitation
Veeramreddy et al. [5]	NSL-KDD	Metaheuristic technique, canonical correlation analysis	FCAAIS minimized the process complexity of designing the scale using FAIS	Less detection accuracy
Hossein [6]	KDD cup 99	K-nearest neighbor (K-NN) and K-means clustering (KMC) algorithms, information theory	Good approach for detecting U2R and probe attacks	Fails to detect "phf" and "mailbomb" attacks
Yang et al. [7]	KDD cup 99, NSL-KDD	New deep neural network (NDNN)	Use of adaptive moment estimator makes the parameter more stable and gives an increased and more effective convergence rate	The increase in network structure will lead to exponential progression of computing time and other issues
Mohammed et al. [8]	KDD cup 99, NSL-KDD and Kyoto 2006 +	Least square support vector machine-based IDS (LSSVM-IDS)	The proposed feature selection method eliminates the redundancy parameter required in MIFS and MMIFS	The overall detection rate is low
Yihan et al. [9]	KDD cup 99	Convolutional neural network (CNN), principal component analysis (PCA), auto-encoder (AE)	The proposed method shares same convolutional kernels, which reduces the number of parameters and calculation amount of training and hence can quickly identify attack type of traffic data	Low detection rates of U2R and R2L

Table 19.2 Features of KDD cup 99

Feature No	Feature Label	Feature No	Feature Label
1	duration	22	is_guest_login
2	protocol_type	23	count
3	service	24	srv_count
4	flag	25	serror_rate
5	src_bytes	26	srv_serror_rate
6	dst_bytes	27	rerror_rate
7	land	28	srv_rerror_rate
8	wrong_fragment	29	same_srv_rate
9	urgent	30	diff_srv_rate
10	hot	31	srv_diff_host_rate
11	num_failed_logins	32	dst_host_count
12	logged_in	33	dst_host_srv_count
13	num_compromised	34	host_same_srv_rate
14	root_shell	35	dst_host_diff_srv_rate
15	su-attempted	36	dst_host_same_src_port_rate
16	num_root	37	dst_host_srv_diff_host_rate
17	num_file_creations	38	dst_host_serror_rate
18	num_shells	39	dst_host_srv_serror_rate
19	num_access_files	40	dst_host_rerror_rate
20	num_outbound_cmds	41	dst_host_srv_rerror_rate
21	is_host_login		

19.4 Proposed Methodology

The workflow of the proposed model is shown in Fig. 19.1. It consists of data collection, preprocessing, feature selection and a classifier. The work is implemented in Weka which is a workbench for machine learning.

The data collected from the Web are first converted into comma-separated value (CSV) format. The class labels are given in different attack names, and since it is anomaly-based detection, all the attacks are assigned as anomaly and remaining as normal.

After preprocessing, the next step is feature selection. "Feature selection methods are used to reduce the dimensionality of the data by removing the redundant and

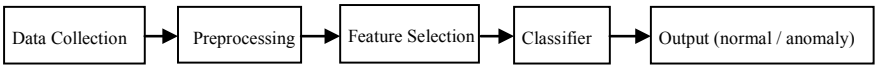


Fig. 19.1 Workflow of the proposed model

irrelevant attributes in a dataset" [10]. In the study, four attributes evaluation, viz. gain ratio, correlation, information gain and symmetrical uncertainty are used. The training data of each dataset run through one attribute evaluator at a time using ranker as the search method which produces some scores for all attributes. And the mean, median and mode of the scores of each attribute evaluator are calculated.

Only those features are selected whose scores are more than or equal to the calculated mean, median and mode separately. Thereafter, selection of features is performed in such a way that the occurrence of the feature is more than or equal to three (since four attribute evaluators are used). Equations (19.1), (19.2) and (19.3) are used for calculating mean, median and mode, respectively.

$$\text{Mean} = \frac{\text{Sum of all feature scores}}{\text{Number of features}} \tag{19.1}$$

$$\text{Median} = \begin{cases} \frac{n}{2}, & \text{when } n \text{ is odd} \\ \frac{(n+1)}{2}, & \text{when } n \text{ is even} \end{cases} \tag{19.2}$$

where n is the number of features

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) \times h \tag{19.3}$$

where l defines the lower limit, h represents the size of the class interval, and f_1, f_0 and f_2 are the frequency of the modal class, class preceding the modal class and class succeeding the modal class, respectively.

In order to get the best features, we have taken the mean and median (as in case of mode, the value comes out to be zero and since the scores of all the features were greater than zero, we need to select all the features and hence it was discarded) as the selection criteria. It was found that the features selected based on median have higher accuracy for both KDD cup 99 and NSL-KDD datasets. These median-based features were used to train the classifiers. The selected features for KDD cup 99 and NSL-KDD datasets are given in Tables 19.3 and19.4, respectively.

The performances can be calculated from the resultant confusion matrix (Table 19.5). For calculating the TP rate and recall, Eq. (19.4) has been used. Equations (19.5), (19.6), (19.7) and (19.8) were used for calculating the FP rate, precision, F-measure and accuracy, respectively.

$$\text{True positive rate (TPR)} = \text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{19.4}$$

$$\text{False positive rate (FPR)} = \text{FP}/(\text{FP} + \text{TN}) \tag{19.5}$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{19.6}$$

Table 19.3 Selected features of KDD cup 99

Feature no.	Feature label	Feature no	Feature label
1	Duration	29	Same_srv_rate
2	Protocol_type	20	Diff_srv_rate
3	Service	31	Srv_diff_host_rate
4	Flag	32	Dst_host_count
5	Src_bytes	34	Host_same_srv_rate
6	Dst_bytes	35	Dst_host_diff_srv_rate
12	Logged_in	36	Dst_host_same_src_port_rate
23	Count	37	Dst_host_srv_diff_host_rate
24	Srv_count	38	Dst_host_serror_rate
25	Serror_rate	39	Dst_host_srv_serror_rate

Table 19.4 Selected features of NSL-KDD

Feature no	Feature label	Feature no	Feature label
3	Service	31	Srv_diff_host_rate
4	Flag	32	Dst_host_count
5	Src_bytes	33	Dst_host_srv_count
6	Dst_bytes	34	Host_same_srv_rate
12	Logged_in	35	Dst_host_diff_srv_rate
23	Count	36	Dst_host_same_src_port_rate
25	Serror_rate	37	Dst_host_srv_diff_host_rate
26	Srv_serror_rate	38	Dst_host_serror_rate
29	Same_srv_rate	39	Dst_host_srv_serror_rate
30	Diff_srv_rate	40	Dst_host_rerror_rate

Table 19.5 Confusion matrix

Predicted			
a	b		
TP	FP	**a = normal**	**Actual**
FN	TN	**b = anomaly**	

$$\text{F} - \text{measure} = 2/\left(1/\text{Precision}\right) + \left(1/\text{Recall}\right) \tag{19.7}$$

$$\text{Accuracy} = \left(\text{TP} + \text{TN}\right)/\left(\text{TP} + \text{TN} + \text{FP} + \text{FN}\right) \tag{19.8}$$

Table 19.6 Algorithms of classifiers

Algorithm for naive bayes [12]	Algorithm for SMO [13]
Step 1: Read the training dataset T; Step 2: Calculate the mean and standard deviation of the predictor variables in each class; Step 3: Repeat Calculate the probability of f_i using the Gauss density equation in each class; Until the probability of all predictor variables (f_1, f_2, f_3,…, f_n) has been calculated Step 4: Calculate the likelihood for each class; Step 5: Get the greatest likelihood;	Step 1: Input C, kernel, kernel parameter and epsilon Step 2: Initialize $\alpha_i \leftarrow 0$ and $b \leftarrow 0$ Step 3: Let $f_a(x) = b + \sum_{i=1}^{m} y_i \alpha_i k(x, x_i)$ and τ the tolerance Step 4: Find Langrange multiplier α_i which violates KKT optimization Step 5: Choose second multiplier and optimize pair. Repeat steps 4 and 5 till convergence Step 6: Update α_1 and α_2 in one step α_1 can be changed to increase $f_a(x_1)$ α_2 can be changed to decrease $f_a(x_2)$ Step 7: Compute new bias weight b
Algorithm for J48 Classifier [14]	**Algorithm for Random Forest** [15]
Step 1: If all cases are of the same class, the tree is a leaf and so the leaf is returned labeled with this class; Step 2: For each attribute, calculate the potential information provided by a test on the attribute (based on the probabilities of each case having a particular value for the attribute). Also calculate the gain in information that would result from a test on the attribute (based on the probabilities of each case with a particular value for the attribute being of a particular class); Step 3: Depending on the current selection criterion, find the best attribute to branch on	Step 1: The algorithm selects random samples from the dataset provided Step 2: The algorithm will create a decision tree for each sample selected. Then it will get a prediction result from each decision tree created Step 3: Voting will then be performed for every predicted result. For a classification problem, it will use mode, and for a regression problem, it will use mean Step 4: And finally, the algorithm will select the most voted prediction result as the final prediction

19.4.1 The Classifiers

"A classifier is any algorithm that sorts data into labeled classes, or categories of information"[11]. There are four classifiers, namely Naive Bayes, sequential minimal optimization, J48 and random forest which are used in the study and are run separately for the comparison of performance. Table 19.6 shows the algorithms of classifiers used.

19.5 Results and Discussion

In KDD cup 99, the Naive Bayes takes the least time with 2.26 s with an accuracy of 98.37% (See Table 19.7). The SMO takes 3249.26 s, 54 min approx. with an accuracy of 99.18%. The J48 takes 22.49 s, giving an accuracy of 99.97%. RF takes

Table 19.7 Performance of KDD cup 99 and NSL-KDD datasets using various classifiers

Dataset	Classifier	Metrics	Time taken (in sec)	TPR	FPR	Precision	F-Measure	Accuracy (%)
KDD cup 99	Naive Bayes		2.26	0.984	0.012	0.985	0.984	98.37
	SMO		3249.26	0.992	0.004	0.992	0.992	99.18
	J48		22.49	1	0	1	1	99.97
	RF		163.24	1	0	1	1	99.98
NSL-KDD	Naive Bayes		0.78	0.904	0.104	0.908	0.903	90.39
	SMO		1196.11	0.973	0.029	0.973	0.973	97.28
	J48		12.21	0.998	0.002	0.998	0.998	99.76
	RF		56.08	0.999	0.001	0.999	0.999	99.91

163.24 s, approximately 2 min, giving the maximum accuracy of 99.98%. If we go on with the amount of accuracy, RF can be chosen as the best classifier among the rest. However, considering the time taken in building the model as another factor along with the accuracy, the J48 can be a better choice.

In NSL-KDD, the Naive Bayes takes 0.78 s, giving an accuracy of 90.39%. SMO takes 1196.11 s approximately equal to 19 min with an accuracy of 97.28%. The J48 takes 12.21 s in building the model achieving an accuracy of 99.76%. The RF takes 56.08 s and gives the highest accuracy of 99.91%. Though SMO yields a strong accuracy rate, it is not preferable since the amount of time it takes in building the model is high. The Naive Bayes cannot be a good option because of its low accuracy rate, though it is a fast-building model. We can opt for J48 or RF according to our needs and requirements about how fast we want to detect and how accurate we want to get. The graph representation of the four learning algorithms based on accuracy is shown in Fig. 19.2.

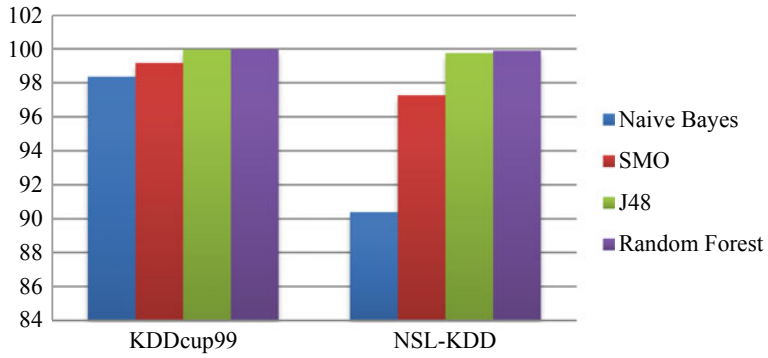


Fig. 19.2 Performance comparison of the four learning algorithms based on accuracy

19.6 Conclusion

From the above analysis, we can see that the accuracy rate based on the KDD cup 99 is greater than the NSL-KDD dataset on every classifier model. This is due to the fact that the KDD cup 99 dataset contains many redundant records which results in the classifier to fall bias toward the more frequent records, hence giving more accuracy rate. The experiment results show that the J48 and random forest classifier effectively detects the intrusion detection using feature reduction methods.

Acknowledgements The authors would like to thank the reviewers whose comments and feedbacks would help in the improvement of the paper.

References

1. Anderson, J.P.: Computer Security Threat Monitoring and Surveillance. James P. Anderson Co. Box 42 Fort Washington, Pa. 19034 (1980)
2. Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J.K.: Network anomaly detection: methods, systems and tools. IEEE Commun. Surv. Tutor. **16**(1), 303–336. First Quarter (2014). https://doi.org/10.1109/SURV.2013.052213.00046
3. Singh, K.J., De, T.: DDOS Attack detection and mitigation technique based on http count and verification using CAPTCHA. In: 2015 International Conference on Computational Intelligence and Networks, Bhubaneshwar, pp. 196–197 (2015). https://doi.org/10.1109/CINE.2015.47
4. Veeramreddy, J., Prasad, V., Koneti, P.: A review of anomaly based intrusion detection systems. Int. J. Comput. Appl. **28**, 26–35 (2011). https://doi.org/10.5120/3399-4730
5. Jyothsna, V., Prasad, K.M.: Anomaly-based intrusion detection system. In: Computer and Network Security, Jaydip Sen, IntechOpen (2019). https://doi.org/10.5772/intechopen.82287
6. Shirazi, H.M.: Anomaly intrusion detection system using information theory, K-NN and KMC algorithms. Austral. J. Basic Appl. Sci. **3**(3), 2581–2597 (2009)
7. Jia, Y., Wang, M., Wang, Y.: Network intrusion detection algorithm based on deep neural network. IET Info. Secur. **13**(1), 48–53 (2018). https://doi.org/10.1049/iet-ifs.2018.5258
8. Ambusaidi, M.A., He, X., Nanda, P., Tan, Z.: Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans. Comput. **65**(10), 2986–2998 (2016). https://doi.org/10.1109/TC.2016.2519914
9. Xiao, Y., Xing, C., Zhang, T., Zhao, Z.: An intrusion detection model based on feature reduction and convolutional neural networks. IEEE Access **7**, 42210–42219 (2019). https://doi.org/10.1109/ACCESS.2019.2904620
10. Al Janabi, K.B.S., Kadhim, R.: Data reduction techniques: a comparative study for attribute selection methods. Int. J. Adv. Comput. Sci. Technol. **8**(1), 1–13 (2018)
11. Johnson Singh, K., De, T.: Efficient classification of DDoS attacks using an ensemble feature selection algorithm. J. Intell. Syst. **29**(1), 71–83 (2020)
12. Saputra, M.F.A., Widiyaningtyas, T., Wibawa, A.: Illiteracy classification using K means-naive bayes algorithm. JOIV : Int. J. Inf. Visual. (2018). https://doi.org/10.30630/joiv.2.3.129
13. Deepa, S.N., Devi, B.A.: Neural networks and SMO based classification for brain tumor. In: 2011 World Congress on Information and Communication Technologies, pp. 1032–1037, Mumbai (2011). https://doi.org/10.1109/WICT.2011.6141390
14. Korting, T.S.: C4.5 Algorithm and multivariate decision trees. Image Processing Division, National Institute for Space Research–INPE (2006)

15. David, D.: Random forest classifier tutorial: how to use tree-based algorithms for machine learning. Free code camp. https://www.freecodecamp.org/news/how-to-use-the-tree-based-algorithm-for-machine-learning

Chapter 20
A Machine Learning Approach to Intrusion Detection System Using UNSW-NB-15 and CICDDoS2019 Datasets



Akoijam Priya Devi and Khundrakpam Johnson Singh

Abstract IDS has gained a lot of importance in the recent years for protecting systems and networks from various kinds of attacks. Unfortunately, as new attacks arise and old attacks are evolving, it cannot detect all kinds of attacks. So there comes a need to optimize the efficiency of detecting an intrusion by choosing best characteristics features and classification methods. In this paper, modern attack-based dataset like UNSW-NB-15 and CICDDoS2019 datasets are used. Main objective of the research work is classification of multi-class attacks for anomaly detection system using machine learning techniques. Using Ranker as a search method and Information Gain, Gain Ratio, One Rule, and Correlation Attribute Evaluators, feature selection is done. Analysis of performance of different machine learning algorithms is also done.

20.1 Introduction

Due to the development in Internet and technology, hacking incidents are increasingly reported by various companies every year. New attack techniques and exploits are continuously being developed to compromise the defense mechanisms. Hence, there is a need for more promising second lines of defense after firewalls. Intrusion detection has become one of the most important cybersecurity measures to be taken to prevent from various kinds of cyberattack. An IDS is a security tool which monitors the network traffic and alerts the system or the network administrator of any suspicious activities.

The thesis overview is listed as Sect. 20.2 discusses the literature review of some of the papers related to IDS. Section 20.3 deals with the datasets. Section 20.4 deals with proposed methodology. Section 20.5 deals with the results and conclusion. Section 20.6 contains conclusion and future works and lastly the references.

A. Priya Devi (✉) · K. Johnson Singh
Department of CSE, National Institute of Technology, Manipur, Imphal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_20

Table 20.1 Literature survey

Authors	Title	Classification	Datasets	Performance
Johnson et al. [1]	Detection and differentiation of application layer DDoS attack from flash events using fuzzy-GA computation	GA-Fuzzy	CAIDA 2007, EPAHTTP, 1998 FIFA World Cup	98.40%
Moustafa et al. [2]	A hybrid feature selection for network intrusion detection systems	EM, NB LR	UNSW-NB-15	Accuracy(LR) 83.0% FAR(EM) 13.1%
Preeti et al. [3]	A detailed investigation and analysis of using machine learning techniques for intrusion detection	ML	KDD'99, UNSW-NB-15	N/A
Moustafa et al. [4]	USWN-NB15: A comprehensive dataset for network intrusion detections systems (UNSW-NB15 network dataset)	ML	UNSW-NB-15	N/A

20.2 Related Works

Table 20.1 given below shows the literature survey on some IDS using machine learning.

Singh et al. [5] proposed a clustering-based model with two clustering methods, namely Gaussian mixture and K-means clustering. The similarity of two different samples is evaluated with distance measure. Shorter the distance between two samples, higher is the similarity. This method gives 99.83% accuracy rate and a FAR of 0.04%.

20.3 Dataset

There are many benchmark datasets [6] available publicly for training and testing of various anomaly detection methods. However, some of these datasets have drawbacks. Datasets like DARPA (1998–1999) and KDD'99 Cup consist of outdated attacks and are less effective on modern networks and infrastructures. Although NSL-KDD dataset removes redundant instances, still some drawbacks of KDD Cup

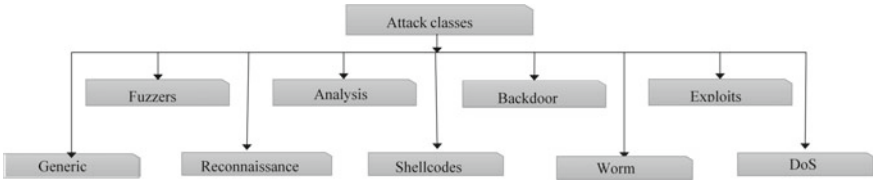


Fig. 20.1 Attack classes of UNSW-NB-15 dataset

remains. CAIDA (2002–2016) dataset lacks traffic feature relevant for identifying individual attacks and also lacks most of the common attacks. UNIBS has only raw traces of traffic and lacks relevant traffic features.

Considering the above drawbacks, we search for a dataset that has current network scenario, less redundant records, one which is properly labeled, and is non-anonymized. Thus, two datasets are considered for my study: UNSW-NB-15 and CICDDoS2019.

20.3.1 UNSW-NB-15 Dataset

The UNSW-NB-15 dataset was created by IXIA Perfect Storm tool in Cyber Range Laboratory of Australian Centre for Cyber Security (ACCS) [7].

Figure 20.1 shows nine attack classes of UNSW-NB-15 dataset. In Fuzzers, an attacker attempts to find the security vulnerabilities in the operating system, program, or in the network due to which resource becomes suspended for some time period and in some case crashing them. Analysis uses combination of various attacks in order to penetrate in a Web site or application. Backdoor are attackers attempt to discover information by bypassing normal authentication process and obtain illegal remote access to the system.

In Exploit, attackers exploits vulnerabilities, glitch or bug in the OS or software, and launch exploits to the system with an intent to cause damage. In Generic, the attacker tries to break the security key of a system. In Reconnaissance, attackers gather information about the victim network to bypass the security control. Shellcode is a code that runs in the victim's machine to gain remote access on the local machine. Worms are malicious programs which can replicate itself to spread to other systems. A DoS attack causes service unavailability to the intended users.

20.3.2 CICDDoS2019 Dataset

CICDDoS2019 dataset [8] contains many new attacks which can be carried out in the application layer. Figure 20.2 shows the output classes of DDoS attack. A DDoS attack is a DoS attack launched from several compromised device that are

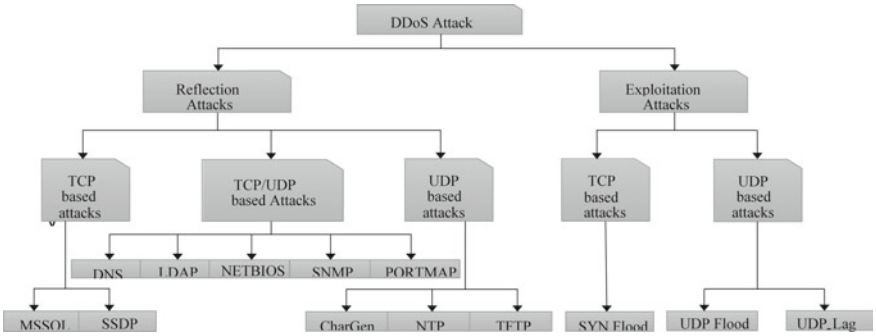


Fig. 20.2 Graphical representation of the output classes of CICDDoS2019 dataset

Table 20.2 Dataset description of UNSW-NB-15 and CICDDoS2019 datasets

Dataset	No. of training records	No. of features	No. of output class
UNSW-NB-15	175,341	49	9
CICDDoS2019	120,848	88	12

distributed globally [9]. Reflection-based DDoS attack can be exploited through the application layer protocols using the TCP, UDP, or combination of both TCP and UDP. MSSQL and SSDP belong to TCP based attack, whereas CharGen, TFTP, and NTP belong to UDP based attacks. Attacks like DNS, NETBIOS, SNMP, and LDAP can be exploited by using either UDP or TCP. Exploitation-based-DDoS attack can be executed through the application layer protocol using TCP and UDP. SYN Flood belong to TCP based exploitation attack category. UDP Flood and UDP_Lag belong to the UDP based exploitation attack category. A dataset description of the two benchmark datasets is given in Table 20.2.

20.4 Proposed Methodology

The workflow diagram of the proposed methodology is shown in Fig. 20.3. It is a model derived from the basic architecture of intrusion detection system [10]. In the first step, the UNSW-NB-15 dataset and CICDDoS2019 dataset are selected. Since the original datasets consists of large amount of data traffic, a part of dataset



Fig. 20.3 Workflow diagram of the proposed model

Table 20.3 UNSW-NB-15 dataset selected features list

Sl	Feature	Sl	Feature name	Sl	Feature name	Sl	Feature name
1	Sbytes	7	Proto	13	ct_dst_sport_ltm	19	Sload
2	Smean	8	ct_srv_dst	14	Sttl	20	Sinpkt
3	Id	9	ct_state_ttl	15	Dload	21	attack_cat
4	Label	10	Dttl	16	ct_src_dport_ltm		
5	Dmean	11	Service	17	ct_dst_src_ltm		
6	Rate	12	ct_srv_src	18	ct_dst_ltm		

Table 20.4 CICDDoS2019 dataset selected feature list

Sl	Feature Name	Sl	Feature Name	Sl	Feature Name
1	Timestamp	12	Source Port	23	Fwd IAT Std
2	Average Packet Size	13	Fwd HeaderLength	24	Subflow Fwd Packets
3	Subflow Fwd Bytes	14	act_data_pkt_fwd	25	Total Fwd Packets
4	Total Length of Fwd Packets	15	Init_Win_bytes_backward	26	Init_Win_bytes_forward
5	Packet length mean	16	Fwd Packets/s	27	ACK Flag Count
6	Avg Fwd Segment Size	17	Flow Packets/s	28	Protocol
7	Fwd Packet Length Mean	18	Flow IAT Mean	29	Min Packet Length
8	Max Packet Length	19	Fwd IAT Total	30	Fwd IAT Max
9	Fwd Packet Length Min	20	Flow IAT Max	31	Flow Duration
10	Fwd Packet Length Max	21	Fwd IAT Mean	32	Flow Bytes
11	Fwd Packet Length Std	22	Flow IAT Std	33	Label

is then extracted and preprocessed in the second step. After using Ranker [11] as a search method and InfoGain, GainRatio, OneR, and Correlation Attribute Evaluators [12], feature selection is done. Twenty-one attributes are selected for UNSW-NB-15 dataset and 33 attributes from the CICDDoS2019 dataset. Tables 20.3 and 20.4 show the selected feature list of UNSW-NB-15 and CICDDoS09 datasets, correspondingly.

20.4.1 J48 Classifier

J48 algorithm is the C4.5 [13] open source implementation in Weka. It accepts discrete and continuous value attributes. Firstly, entropy measure is computed for

each feature. Then, set of examples are partitioned in accordance to the possible outcome values of those features which has the lowest entropy. Probabilities are then estimated.

20.4.2 Naïve Bayes Classifier

NB is based on the Bayes' theorem as shown in Eq. (20.1), where posterior probability of class c is calculated given a predictor x.

$$P(c/x) = (P(x/c) \times P(c))/(P(x)) \tag{20.1}$$

Given a training dataset, first convert it into a frequency table. Next, create a likelihood table by computing the probabilities of each feature. Using naïve Bayesain equation, the posterior probability of each of the class are computed. The class having the highest posterior probability becomes the prediction output.

20.4.3 Decision Table Classifier

A decision table (DT) has a more ordered set of "if–then" rules which is more compact, understandable, and simpler than a decision tree [14]. It is usually done by the tables' cross-validation performance measurement for all the different attributes' subsets and then chose the subset with the best performance.

20.4.4 Simple Logistic Based Classifier

Simple logistic regression (SL) is comparable to the linear regression except in the fact that the dependent variable is considered as nominal and not as a measurement. It finds out the equation which predicts best value of a variable Y for each variable X value.

20.5 Results and Discussions

Generation of a performance metrics of a model is a crucial part in understanding the performance of a classification model. Table 20.5 shows the TP Rate, FP Rate, Precision, Recall, F-Measure, and ROC Area [15] of each of the output classes of UNSW-NB-15 using J48, NB, DT, and SL. Table 20.6 gives the overall performance

Table 20.5 Performance metrics for each classes of UNSW-NB-15 using J48, NB, DT, and SL

	Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
Backdoor	J48	0.241	0.002	0.577	0.241	0.340	0.857
	NB	0.356	0.066	0.051	0.356	0.090	0.863
	DT	0.088	0.000	0.705	0.088	0.156	0.832
	SL	0.006	0.000	0.688	0.006	0.012	0.895
Analysis	J48	0.241	0.002	0.577	0.241	0.327	0.872
	NB	0.168	0.011	0.146	0.168	0.156	0.892
	DT	0.123	0.001	0.704	0.123	0.209	0.866
	SL	0.014	0.000	0.903	0.014	0.028	0.931
Fuzzers	J48	0.887	0.005	0.953	0.887	0.919	0.974
	NB	0.283	0.035	0.483	0.283	0.357	0.877
	DT	0.823	0.010	0.902	0.823	0.861	0.977
	SL	0.796	0.037	0.716	0.796	0.754	0.967
Shellcode	J48	0.742	0.002	0.712	0.74	0.727	0.938
	NB	0.907	0.128	0.044	0.907	0.084	0.939
	DT	0.542	0.002	0.597	0.542	0.568	0.976
	SL	0.000	0.019	0.000	0.000	0.000	0.951
Reconnaisance	J48	0.757	0.005	0.906	0.757	0.825	0.963
	NB	0.225	0.138	0.094	0.225	0.133	0.800
	DT	0.759	0.005	0.915	0.759	0.829	0.967
	SL	0.561	0.017	0.675	0.561	0.613	0.952
Exploit	J48	0.852	0.096	0.675	0.85	0.753	0.943
	NB	0.379	0.016	0.851	0.379	0.524	0.839
	DT	0.898	0.136	0.608	0.898	0.725	0.942
	SL	0.863	0.135	0.600	0.863	0.708	0.942
DoS	J48	0.298	0.034	0.394	0.298	0.340	0.901
	NB	0.004	0.001	0.242	0.004	0.007	0.861
	DT	0.041	0.003	0.475	0.041	0.076	0.916
	SL	0.081	0.010	0.375	0.081	0.134	0.913
Worms	J48	0.531	0.000	0.627	0.531	0.575	0.924
	NB	0.177	0.011	0.012	0.177	0.023	0.911
	DT	0.508	0.000	0.574	0.508	0.539	0.958
	SL	0.000	0.000	0.000	0.000	0.000	0.975
Generic	J48	0.985	0.001	0.995	0.985	0.990	0.996
	NB	0.973	0.008	0.974	0.973	0.974	0.989
	DT	0.979	0.000	0.998	0.979	0.989	0.995
	SL	0.978	0.002	0.992	0.978	0.985	0.998

Table 20.6 Overall performance metrics for UNSW-NB-15 dataset

	Accuracy (%)	Kappa statistic	MAE	RMSE	RAE	RSE (%)
J48	87.45	0.8399	0.0307	0.1311	19.405%	46.6264
NB	60.2323	0.5227	0.0801	0.2642	50.6788	93.9938
DT	85.3485	0.811	0.0539	0.1502	34.1321%	53.4351
SL	81.1231	0.7592	0.0512	0.1635	32.3922%	58.1529

metrics for UNSW-NB-15 by calculating their Accuracy, Kappa statistic, Mean absolute error (MAE), Root mean squared error (RMSE), Relative absolute error (RAE), and Relative squared error (RSE) using the classifier algorithms. Table 20.7 compares the performance metrics for each output classes for CICDDoS2019 using J48, DT, NB, and SL. Table 20.8 gives the overall performance metrics of CICDDoS2019.

Figure 20.4 shows a chart table comparing the accuracy of the two datasets w.r.t. each classifiers used. It is observed that the classifiers have better performance in case of CICDDoS2019 dataset.

20.6 Conclusions and Future Scope

Comparison of performance metrics is done for the two benchmark datasets. J48 gives highest accuracy of 87.45% for UNSW-NB-15 and 99.65% for CICDDoS2019. DT gives 60.23% for UNSW-NB-15 and 99.08% for CICDDoS2019. NB gives 85.35%

Table 20.7 Performance metrics for each classes of CICDDoS2019 using J48, DT, SL, and NB

	Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC
DrDoS_ DNS	J48	0.982	0.001	0.995	0.982	0.988	0.999
	DT	0.978	0.004	0.713	0.978	0.975	1.000
	SL	0.968	0.004	0.971	0.968	0.970	0.984
	NB	0.793	0.012	0.900	0.789	0.846	0.976
Benign	J48	0.956	0.000	0.994	0.968	0.981	0.988
	DT	0.728	0.001	0.941	0.728	0.821	0.993
	SL	0.983	0.000	0.986	0.983	0.985	0.992
	NB	0.991	0.003	0.839	0.991	0.909	1.000
DrDoS_ LDAP	J48	1.000	0.001	0.995	1.000	0.997	1.000
	DT	1.000	0.000	1.000	1.000	1.000	1.000

(continued)

Table 20.7 (continued)

	Classifier	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC
	SL	0.985	0.007	0.945	0.985	0.965	0.991
	NB	0.956	0.047	0.724	0.956	0.824	0.988
DrDoS_ MSSQL	J48	1.000	0.001	0.996	1.000	0.998	1.000
	DT	1.000	0.006	0.977	1.000	0.988	1.000
	SL	0.982	0.017	0.933	0.982	0.957	0.984
	NB	0.847	0.016	0.928	0.847	0.886	0.987
DrDoS_ NetBIOS	J48	0.999	0.000	0.998	0.999	0.998	0.999
	DT	1.000	0.000	0.998	1.000	0.999	1.000
	SL	0.986	0.000	1.000	0.986	0.993	0.994
	NB	0.923	0.000	0.997	0.923	0.959	0.997
DrDoS_ NTP	J48	0.999	0.001	0.974	0.987	0.980	0.998
	DT	0.929	0.000	0.996	0.929	0.962	0.999
	SL	0.987	0.000	0.989	0.987	0.988	0.994
	NB	0.892	0.010	0.764	0.892	0.823	0.998
DrDoS_ SSDP	J48	1.000	0.000	1.000	1.000	1.000	1.000
	DT	1.000	0.000	0.999	1.000	1.000	1.000
	SL	0.984	0.000	0.999	0.984	0.992	0.993
	NB	0.866	0.008	0.921	0.866	0.892	0.983
DrDoS_ UDP	J48	0.999	0.000	1.000	0.999	0.999	1.000
	DT	1.000	0.000	1.000	1.000	1.000	1.000
	SL	0.982	0.004	0.968	0.982	0.975	0.990
	NB	0.943	0.017	0.870	0.943	0.905	0.995
Syn	J48	1.000	0.000	1.000	1.000	1.000	1.000
	DT	1.000	0.000	1.000	1.000	1.000	1.000
	SL	0.903	0.003	0.954	0.903	0.928	0.959
	NB	0.996	0.009	0.887	0.996	0.939	1.000
TFTP	J48	0.999	0.000	1.000	0.999	0.999	1.000
	DT	1.000	0.000	1.000	1.000	1.000	1.000
	SL	0.970	0.000	1.000	0.970	0.984	0.987
	NB	0.861	0.001	0.992	0.861	0.922	1.000
UDP-lag	J48	1.000	0.000	0.997	1.000	0.9998	1.000
	DT	1.000	0.000	1.000	1.000	1.000	1.000
	SL	0.910	0.000	1.000	0.910	0.953	0.961
	NB	0.837	0.002	0.971	0.837	0.899	0.982
WebDDoS	J48	0.000	0.000	0.000	0.000	0.000	0.500
	DT	0.000	0.000	0.000	0.991	0.000	0.514
	SL	0.000	0.000	0.000	0.000	0.000	0.647
	NB	1.000	0.003	0.003	1.000	0.007	0.998

Table 20.8 Overall performance metrics of CICDDoS2019 using the classifier models

	Accuracy	Kappa statistic	MAE	RMSE	RAE	RSE
J48	99.6475%	0.996	0.0007	0.0209	0.04732%	7.6939%
NB	88.6742%	0.8725	0.0196	0.1332	13.2728%	49.0389%
DT	99.084%	0.9896	0.0084	0.00447	5.6857%	16.4605%
SL	96.9284%	0.9653	0.0051	0.0715	3.4808%	26.3234%

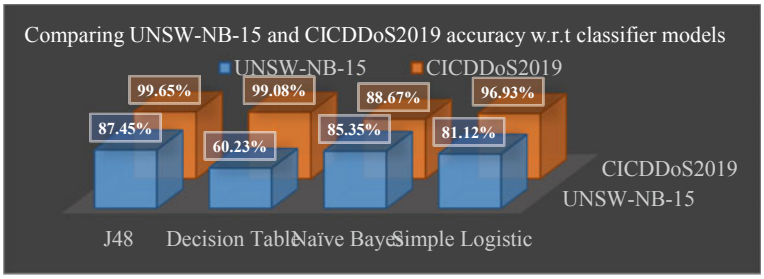


Fig. 20.4 Accuracy of the two datasets with respect to the classifier models

for UNW-NB-15 and 88.67% for CICDDoS2019. SL gives 81.12% for UNSW-NB-15 and 96.93% for CICDDoS2019. This research work also compares the performance metrics of each class such that for a given class output, which algorithm gives the highest score.

Future research will be on in-depth analysis of intrusion detection system models. Hybrid machine learning techniques or deep neural network may be used.

References

1. Singh, K.J., Thongam, K., De, T.: Detection and differentiation of application layer DDoS attack from flash events using fuzzy-GA computation. IET Info. Secur. **12**(6), 502–512 11 (2018)
2. Moustafa, N., Slay, J.: A hybrid feature selection for network intrusion detection systems: central points. In: Proceedings 16th Australian Information Warfare Conference 2015, pp. 1–10. (2015)
3. Mishra, P., Varadharajan, V., Tupakula, U., Pilli, E.S.: A detailed investigation and analysis of using machine learning techniques for intrusion detection. IEEE Commun. Surv. Tutor. **21**(1), 686–728, Firstquarter (2019). https://doi.org/10.1109/COMST.2018.2847722
4. Moustafa, N., Slay, J.: USWN-NB15: a comprehensive data set for network intrusion detections systems (UNSW-NB15 Network Data Set). School of Engineering and Information Technology, Canberra, Australia (2015)
5. Singh, P., Venkatesan, M.: Hybrid Approach for intrusion detection system. In: 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, pp. 1–5 (2018)

6. Bhuyan, M.H., Bhattacharyya, D.K., Kalita, J.K.: Network anomaly detection: methods, systems and tools. IEEE Commun. Surv. Tutor. **16**(1), 303–336, First Quarter (2014)
7. The UNSW-NB15 dataset Available at: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/cybersecurity/ADFA-NB15-Datasets/ (Accessed on 12 Feb 2020)
8. CICDDoS2019 dataset Available at: https://www.unb.ca/cic/datasets/ddos-2019.html (Accessed on 25 January 2020)
9. Johnson Singh, K., Thongam, K., De, T.: Entropy-based application layer DDoS attack detection using artificial neural networks. Entropy **18**, 350 (2016)
10. Heady, R., Luger, G., Maccabe, A., Servilla, M.: The architecture of a network level intrusion detection system. Tech. Rep., Computer Science Department, University of New Mexico, New Mexico (1990)
11. Blum, A.I., Langley, P.: Selection of relevant features and examples in machine learning. Artif. Intell. **97**, 245–271 (1997)
12. Mukherjee, S., Sharma, N.: Intrusion detection using naive bayes classifier with feature reduction. Proc. Technol. **4**, 119–128 (2012)
13. Quinlan, J.R.: In: C4.5: Programs for Machine Learning, San Mateo, Morgan Kaufman (1993)
14. Kalmegh, S.R.: Comparative analysis of the WEKA classifiers rules conjunctiverule and decisiontable on indian news dataset by using different test mode. Department of Computer Science, Sant Gadge Baba Amarvati University, Amaravati(M.S.), India (2018)
15. Shone, N., Ngoc, T.N., Phai, V.D., Shi, Q.: A deep learning approach to network intrusion detection. IEEE Trans. Emerg. Topics Comput. Intell. **2**(1), 41–50 (2018)

Chapter 21
Discovery of Localized Malicious Attack in Wireless Networks



S. Gowri, Senduru Srinivasulu, J. Jabez, J. S. Vimali, and A. Sivasangari

Abstract Wireless communication is merely a boon to mankind. With wireless communication, data is transferred between the nodes at a very high rate in recent days. With emerging 5G technology, it is estimated that the devices are going to receive a bandwidth which is as high as 1GBPS which is the usual rate delivered by broadband services. All these are possible with packet data transfer. Packet transfer occurs in wireless networks in two ways which are flooding and routing. However, one challenge in communication is security. We need to ensure that our packets are sent to the receiver securely without any third-party access to our data packets. We propose a system which makes use of the spatial information property of a node and tries to calculate received signal strength (RSS), which is the main parameter for plotting the attacker nodes in the network and eliminating the attacker nodes by using clustering techniques against a radar grid.

21.1 Introduction

Spoofing attack is the most common attack which the attackers use to gain authorization to the devices in a vulnerable way, such that the attacker nodes exist in the system as if they have been authenticated and recognized by the administrator who manages the devices connected on the same network on the receiver side. The most important thing the user of the service thinks about is his data. In a system which does not authenticate its users in a proper way, the data is never secure. If the data is not secure, the usability of the system by its users will be reduced in a significant way. In broad sense, spoofing is nothing but a technique used to gain access to the computer in a vulnerable way by imitating its identification with the available nodes which have already been authenticated by the administrator. Vulnerabilities are the main reason which enable the attackers to enter the system in a vulnerable way. In wired communication systems, the data is kept secure free from vulnerabilities by

S. Gowri (✉) · S. Srinivasulu · J. Jabez · J. S. Vimali · A. Sivasangari
Sathyabama Institute of Science and Technology Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai 600 119, Tamilnadu, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_21

using certain selected encryption algorithms by using several techniques which are sole part of crypt arithmetic which comes under a branch of cryptography whose main motive is to encrypt the data. However, historical studies prove that 90% of the times wired communication is very secure where it will be quite cumbersome for the attackers to enter the system vulnerably.

At present, we have several technologies which run on client side to ensure security. This attack brings us to the most popular point of using API's in computing where we do not have access to the database of the authenticator, and instead, we have several application programming interfaces which enable us to hit them and retrieve the data in the most secured way by generating authentication tokens. The concepts on which this token authentication relies on is BASICAUTH, APIKEY, OAUTH1, and OAUTH2 whose primary purpose is authentication in a more secured and a controlled way. Providing access to the data in a controlled way is quite a decent job to be done before transmitting the data through a network. This can be done by using access specifiers and access modifiers which secure data by acting as their companions. Several strategies have been proposed promote security for identifying and removing the spoofing attackers [1]. One such strategy is crypt arithmetic where we generate private key and send it as a secret key to the receiver [2, 3].

21.2 Literature Survey

One instance of computer security is wireless security, which is nothing but the collection of security threats and security concerns which are most prevalent in wireless networks. Organizations that follow a local area wireless network are specifically concerned about the vulnerabilities to their security breaches which are caused by rogue access points. These countermeasures secure both the network and the valuable materials that are held by the network. But there is one restriction for applying such actions, and the restriction is that it has to be applied to the network devices uniformly. Wireless communication technologies have been increased in domains beyond the original intended, original areas due to its availability, and low cost. Several industrial applications have security requirements which are pretty specific to the application under consideration.

Hence, industrial application characteristics are important to understand, and the vulnerabilities have to be evaluated bearing the highest risk in the context. Wireless networks are easy to break, so the hackers are looking forward to breaking the wireless networks [4]. Usability of wireless networks has grown enormously to an extent that every individual or any enterprise relies mostly on this type of communication to accomplish most of its networking services. A spoofing attack [5, 6] is the most common attack which the attackers use to gain authorization to the devices in a vulnerable way, such that the attacker nodes exist in the system as if they have been authenticated and recognized by the administrator who manages the devices connected on the same network on the receiver side. Vulnerabilities are the main reason which enables the attackers to enter the system in a vulnerable way. With

current 802.11 protection strategies, like wired equivalent privacy (WEP), Wi-Fi Safe Access (WPA), or 802.11i, such methods can only secure data frames. Spoofing attacks can further enable a range of traffic intrusion attacks such as access control list attacks, rogue access point (AP) attacks, and ultimately denial of service (DoS) attacks [7, 8]. Multiple opponents can mask the same identity in a large network and cooperate to conduct malicious attacks such as attacks on network resources and denial of service.

21.3 Proposed Methodology

Methodology is quietly the routine model view controller (MVC) where we have several view controllers and they are modeled in such a way that they follow a schematic representation where on procedure calls another and the other call the later and it goes on until the entire system has performed the tasks assigned to them. Our proposed system is going to use a specific property which is known as spatial property which holds the information of a particular node which is concerned with its physical identification and uniqueness which are hard to imitate. Later, it turns out that this specific property makes the node potentially secure by overthrowing the hackers who tries to imitate that node. Usage of this property is often considered as a privilege because the cost of state of the architecture is very low because all these properties associated with the node comes as a package without ourselves having to explicitly provide the functionality to the network nodes. By this way, we are actually implementing model which allows us to not only detect the node hackers but also provide us with an opportunity to eliminate them and swipe them out completely from the existing network by showing the number of hacker nodes which tried to enter.

The architecture which we follow is quite simple where we have a server for management, a database for storing information, and users who are authenticated by the server or an outsourcing community. When the users try to access information from the database by communicating with the server, there might be several spoofing attackers who try to enter into the system by keeping track of vulnerabilities which are associated with the existing system. So, we use several algorithms to identify them by plotting and removing them from the system. This might be enough as of now to recognize and remove the attackers from the network and to promote security. Later, it turns out that this specific property makes the node potentially secure by overthrowing the hackers who tries to imitate that node (Fig. 21.1).

The traditional method to manage hinder disparaging ambushes is to use cryptographic-based approval. We have introduced a secure and proficient key management (SEKM) framework [9]. SEKM creates a public key infrastructure (PKI) by applying a secret sharing arrangement and a basic multicast server gathering. We actualized a key management instrument with irregular key resuscitate and host disavowal to prevent the compromise of affirmation keys. An affirmation framework for different leveled, extraordinarily designated sensor frameworks is proposed.

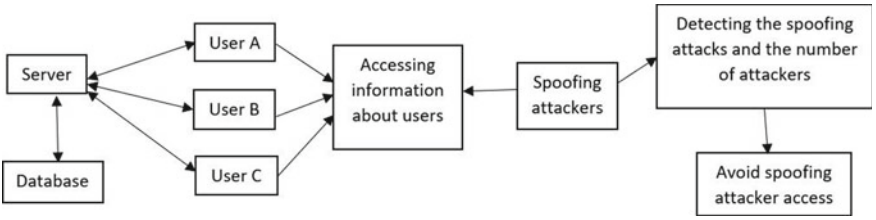


Fig. 21.1 System architecture

In any case, the crypto-sensible approval may not be reliably relevant considering the compelled resources on far off devices and lacking of a fixed key management structure in the distant frameworks.

21.3.1 Detection Model of Generalized Attack

In this segment, we portray our detection model of generalized attack, which comprises of two stages: assault discovery, which recognizes the nearness of an assault, and number assurance, which decides the quantity of foes. The number assurance stage will be introduced.

Framework

The customary restriction approaches are found based on the middle value of RSS from every hub personality contribution to evaluate the situation of a hub. Be that as it may, in remote mocking assaults, the RSS stream of a hub character might be blended in with RSS readings of both the first hub just as satirizing hubs from various physical areas. The conventional strategy for averaging RSS readings cannot separate RSS readings from various areas and consequently is not practical for confining foes. Not quite the same as conventional confinement draws near, our incorporated recognition and limitation framework uses the RSS medoids came back from SILENCE as contributions to restriction calculations to evaluate the places of enemies. The arrival positions from our framework incorporates the area gage of the first hub and the aggressors in the physical space. Taking care of foes utilizing diverse transmission power levels, a foe may change the transmission power levels when performing satirizing assaults with the goal that the limitation framework cannot assess its area precisely. We look at the complex rule models that theys got influence as an element of the separation to the milestone speaks to the transmitting influence of a hub at the reference separation d0, d is the separation between the transmitting hub and the milestone, and is the way misfortune type. Further, we can communicate the distinction of the got power between two tourist spots, I and j, as based on which we found that the distinction of the relating got power between two unique milestones which is free of the transmission power levels. Along these lines, when an enemy dwelling at a physical area fluctuates its transmission capacity to play out a parodying assault,

the distinction of the RSS readings between two unique tourist spots from the foe is consistent since the RSS readings are acquired from a solitary physical area. We would then be able to use the distinction of the medoids vectors in signal space got from SILENCE to restrict foes.

Algorithms *Radar-gridded module.* The radar-gridded calculation is a scene-matching restriction calculation. Radar-gridded utilizes an inserted signal guide, which is worked from a lot arrived at the midpoint of RSS readings with known areas. Given a watched RSS perusing with an obscure area, radar restores the x, y of the closest neighbor in the sign guide to the one to limit, where "closest" is characterized as the euclidean separation of RSS that focuses in a N-dimensional sign space, where N is the quantity of milestones.

Area-based probability module. ABP additionally uses an introduced signal guide. Further, the test zone is separated into an ordinary lattice of equivalent estimated tiles. ABP accepts the conveyance of RSS for every milestone and follows a Gaussian appropriation with mean as the normal estimation of RSS perusing vector s. ABP then registers the likelihood of the remote gadget being at each tile Li, with I ¼ 1...L…, so on.

Bayesian networks. BN limitation is a multi-emphasis calculation that encodes the sign to separate proliferation model into the Bayesian graphical model for confinement of the essential Bayesian network utilized for our examination. The vertices X and Y speak to area; the vertex is the RSS perusing from the milestone; and the vertex Di speaks to the euclidean separation between the area indicated by X and Y and the milestone. The estimation follows a sign proliferation model. Foes utilized a similar transmission power levels as the first hub and the returned medoids are utilized. Enemies changed their transmission power level from 15 to 10 dB, and the returned medoids are utilized; and foes changed their transmission power level from 15 to 10 dB, and the distinction of returned medoids are utilized. The key perception from the presentation of utilizing the distinction of returned medoids in dealing with enemies utilizing diverse transmission power levels is similar to the outcomes when foes utilized a similar transmission power levels as the first hub. Further, the confinement execution is a lot of more regrettable than the customary methodologies if the distinction of returned medoids is not utilized while restricting foes utilizing diverse transmission power levels, appeared as the case 2 above. Specifically, when utilizing our methodology, we can achieve the middle mistake of 13 feet for both radar-gridded and ABP on the off chance that 3, a 40–50% execution improvement, contrasting with the middle blunders of 20 and 19 feet for radar-gridded and ABP, individually, on the off chance that 2. Accordingly, it is exceptionally compelling in restricting various foes with or without changing their transmission power levels.

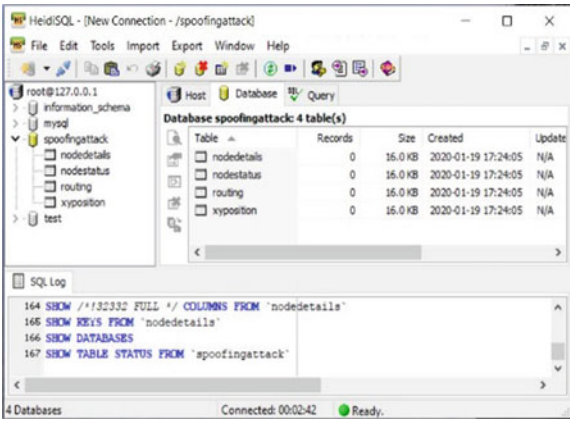


Fig. 21.2 Database connection

21.4 Results and Discussion

A secured network is free from spoofing attackers and unnecessary traffic which enables the network to transmit the data in a safe, secure, and faster way. Even the efficiency and reliability of the network is enhanced as a result of removing unnecessary traffic. Vulnerabilities are the main reason which enable the attackers to enter the system in a vulnerable way. In wired communication systems, the data is kept secure free from vulnerabilities by using certain selected encryption algorithms by using several techniques which are sole part of crypt arithmetic which comes under a branch of cryptography whose main motive is to encrypt the data. However, historical studies prove that 90% of the times wired communication is very secure where it will be quite cumbersome for the attackers to enter the system vulnerably. Wireless networks have grown enormously to an extent that every individual or any enterprise relies mostly on this type of communication to accomplish most of its networking services.

21.5 Screenshots

See Figs. 21.2, 21.3, 21.4, 21.5, 21.6, 21.7, 21.8, 21.9 and 21.10.

21.6 Conclusion and Future Work

Usability of wireless networks has grown enormously to an extent that every individual or any enterprise relies mostly on this type of communication to accomplish most of its networking services. Therefore, it is highly important to promote

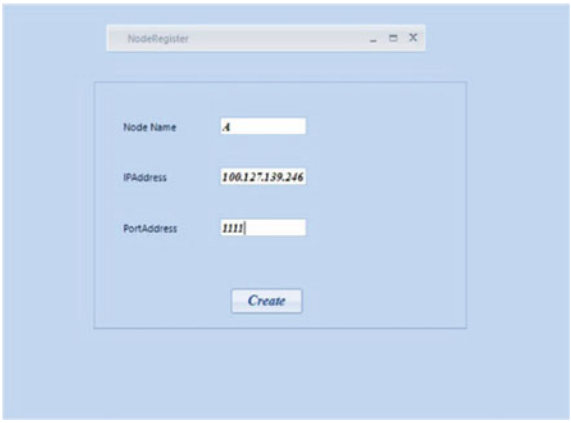


Fig. 21.3 Node creation

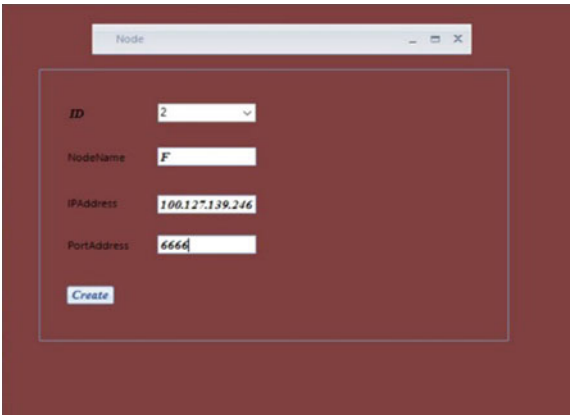


Fig. 21.4 Hacker node

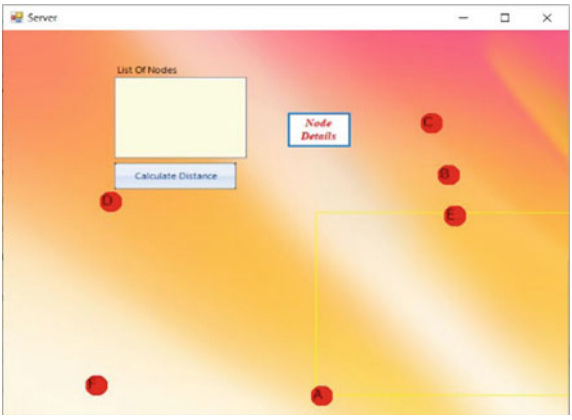


Fig. 21.5 Node details and euclidean dist. calc

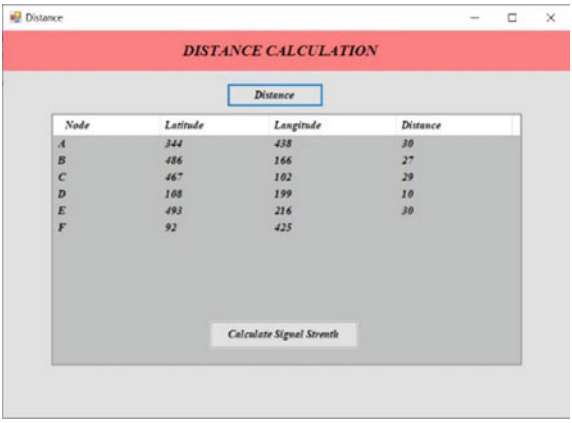


Fig. 21.6 Calculating RSS

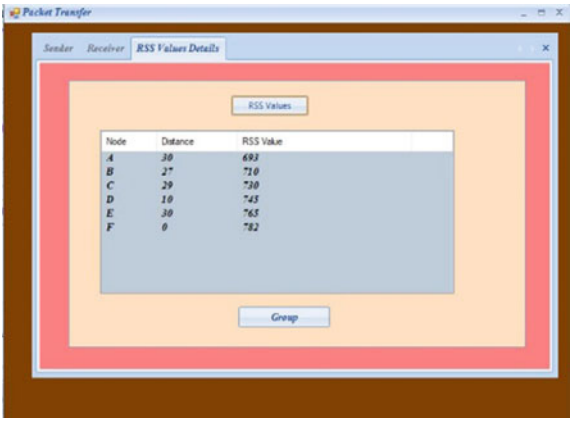


Fig. 21.7 RSS value

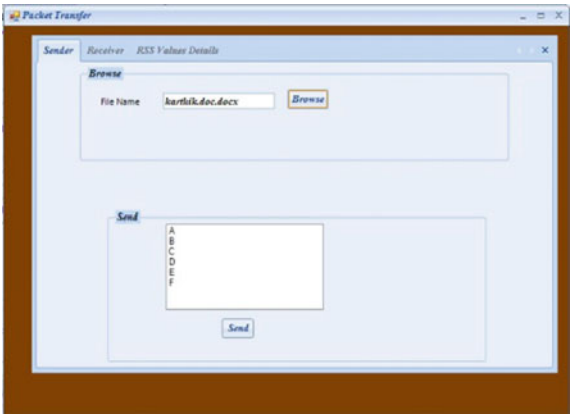


Fig. 21.8 Sending data

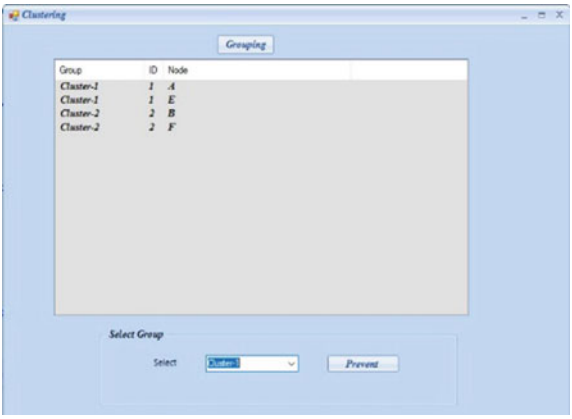


Fig. 21.9 Selecting cluster to prevent attack

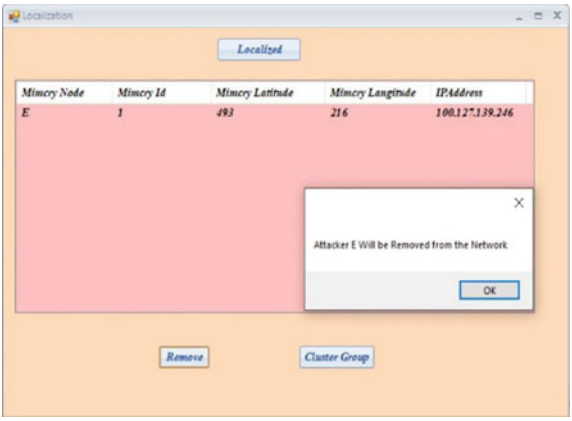


Fig. 21.10 Remove hacker node in network

security, flexibility, consistency, reliability in the wireless networks. As a token of development, it is our responsibility to do enough for promoting this security by identifying the attackers and removing them from the networks. It is important to know the properties associated with a network node because it is the primary entity within the network system to do any kind of operations on the network. We found that our location components are exceptionally powerful in both recognizing the nearness of assaults with recognition rates more than 98% and deciding the quantity of foes, achieving more than 90% hit rates and accuracy all the while when utilizing silence and SVM-based instrument. Further, based on the quantity of assailants controlled by our instruments, our incorporated recognition and confinement framework can restrict any number of foes in any event, when aggressors utilizing distinctive transmission power levels.

References

1. Li, Q, Trappe, W: Relationship-based detection of spoofing- related anomalous traffic in ad hoc networks. In: Proceedings Annual IEEE Communications Society on IEEE and Sensor and Ad Hoc Communications and Networks (SECON), (2006)
2. Wool, A.: Lightweight key management for IEEE 802.11 wireless lans with key refresh and host revocation. ACM/Springer Wireless Netw. **11**(6), 677–686 (2005)
3. Sheng, Y., Tan, K., Chen, G., Kotz, D., Campbell, D.: Detecting 802.11 MAC layerspoofing using received signal strength. In: Proceedings IEEE INFOCOM, April (2008)
4. Faria, D., Cheriton, D.: Detecting Identity-Based Attacks in Wireless Networks Using Signalprints. Proc, ACM Workshop Wireless Security (WiSe) (2006)
5. Yang, J., Chen, Y., Trappe, W.: Detecting spoofing attacks in mobile wireless environments. In: Proceedings Annual IEEE Communications Society Conference Sensor, Mesh and Ad Hoc Comm. And Networks (SECON), (2009)
6. Chen, Y., Trappe, W., Martin, R.P.: Detecting and localizing wireless spoofing attacks. In: Proceedings Annual IEEE Communications Society Conference Sensor Mesh and Ad Hoc Communications and Networks (SECON), May (2007)
7. Bellardo, J., Savage, S.: 802.11 Denial-of-service attacks: real vulnerabilities and practical solutions. In: Proceedings USENIX Security Symposium, pp. 15–28. (2003)
8. Ferreri, F., Bernaschi, M., Valcamonici, L.: Access points vulnerabilities to dos attacks in 802.11 networks. In: Proceedings IEEE Wireless Communications and Networking Conference (2004)
9. Wu, B., Wu, J., Fernandez, E., Magliveras, S.: Secure and efficient key management in mobile ad hoc networks. In: Proceedings IEEE Int'l Parallel and Distributed Processing Symposium (IPDPS) (2005)

Chapter 22
An Adapted Approach of Image Steganography Using Pixel Mutation and Bit Augmentation



Ravi Saini, Kamaldeep Joshi, and Rainu Nandal

Abstract This paper presents a new technique of data hiding in spatial domain using digital image as cover file. It uses two logical modules, binary converter and bit extractor in its processing. Binary converter module converts the identified pixel value into binary sequence of eight bits. Bit extractor module extracts the seven bits of pixel value except LSB. The proposed method hides the message in the seven bits of pixel value except LSB as LSB is the most sensitive bit and can be altered on the network by some error. It disburses the message evenly on the pixel intensity. The preliminary results show that it expels various anomalies associated with some state-of-the-art methods. It also provides better imperceptibility and hiding capacity.

22.1 Introduction

The term steganography is consolidation of two Greek words "Steganos" and "Graph". Steganos means "hidden or covered" and graph means "to write" [1]. So, we can say that steganaography means hidden or covered writing. Steganography is not the new field; it has been used from long interval of time. Its origin has been found in the history at about 440 Before Christ in the book "The Histories" written by Herodotus who was a Greek historian. He was known as "Father of the History". The book "The Histories" contained the detail of the series wars between Greek and Persia from 499 BC to 449 BC. It contains two examples of steganography. Histiaeus was the son of Greek King Lysagoras. His subordinate Aristagoras was in prison. Histiaeus sent a message to Aristagoras by using Steganography. He called his most loyal slave. He trimmed the head of his captive and writes the epistle on the scalp of his servant. He kept his slave with himself until his hair has regrown. After that, he sent his slave to prison to meet Aristagoras. Arsitagoras shaved the head of slave and read the message written on the scalp of slave. The second example

R. Saini (✉)
GCW Gurawara, Rewari, India

K. Joshi · R. Nandal
UIET, Maharshi Dayanand University, Rohtak, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_22

of steganography was found from the side of Persia. Persia King Demaratus sent a message to his subordinates by writing it to wax tablet. Wax tablet is the tablet made up of wood and covered with wax. It provides reusable writing surface. Another example of steganogrpahy was given by Johannes Trithemius who developed Ave-Maria-Cipher. By using this cipher, information can be hiden within Latin praise of God. Invisible ink has also been used as media for secret communication in the history. The message written by invisible ink can be recouped at the receiver end by putting the media in the light or by some other means [2, 3]. The example of steganography is given below: Suppose the following message is transferred from A to B. SIVANSH AND ADMESH DINED HAPPILY. If B holds the second letter of each word of the above message, then he will get the secret message "INDIA". This type of steganography is called null cipher steganography. In this era of digital media, many steganography techniques have been designed and developed so far. These steganography techniques use digital media like text file, image file, audio file, video file, etc., as cover media for secret information transfer [4]. We can divide the steganography techniques into basically four types reckoning on the type of cover media. Text steganography, image steganography, audio steganography, and video steganography are categories of steganography. Image steganography is very popular as compared to other types of steganography techniques in the era of digital world. Various tools have also been developed which provide insertion and retrieval of message automatically [5]. The examples of steganography tools are Steg, STools, Stegmail, Openpuff, Openstego, Outguess, QuickStego, StegoShare, StegFS, Image-Spyer G2, etc. Many image steganography techniques have been developed in recent times like LSB method [6], parity checker method [7], GLM method [8], etc. In this paper, we will propose a new method for image steganography which will expel the anomalies attached with some traditional methods.

22.2 Proposed Method

A new method of image steganography using pixel mutation and bit augmentation is proposed in this section. It uses the any image file as cover file for concealing the secret data. There are two logical modules, binary converter and bit extractor which are used by this approach. At first, the pixel location is selected by using random number generator function where the message bit is to be inserted. The secret message is also translated into bit sequence using the binary converter module. Every selected pixel can be used for hiding only one bit of secret data. The selected pixel is also converted into binary sequence of eight bits using binary converter. The first seven bits of binary pixel value except LSB are extracted using bit extractor module. The message bit is distributed into these seven bits on the basis of insertion process. The insertion process is shown by Fig. 22.1. Sender engenders the stego image at the sender side and sent it to the accepter. The accepter generates the same locations of stego image where message bits are hidden and extract the pixel value of location.

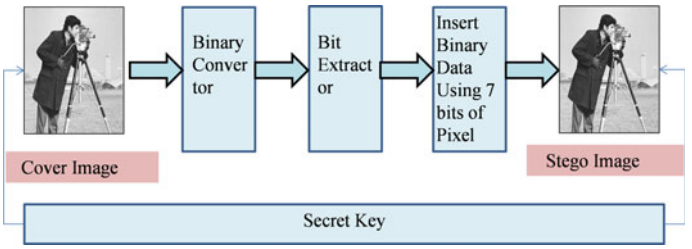


Fig. 22.1 Insertion process

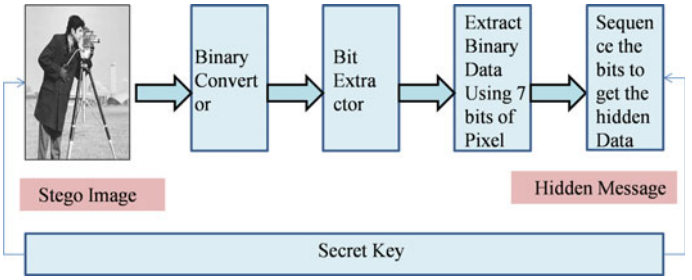


Fig. 22.2 Retrieval process

The message bits are extracted by reverse process at the receiver end and sequenced to form the required message. Secret key is experienced between sender and receiver. The extraction process is shown by Fig. 22.2.

22.3 Algorithm

In this section, we will discuss about the insertion and retrieval algorithm.

22.3.1 Insertion Algorithm

1. Select the cover image in which you want to insert the secret data. Let the cover image is $C(r, c)$ where r symbolizes total rows in the cover image and c symbolizes total columns in the cover image.
2. Convert the secret message M into binary sequence of bits using binary converter.
3. Generate the random location of the pixels using to the Eq. 22.1:

$$Loc(n+1) = (x * Loc(n) + y) Mod\ M \tag{22.1}$$

$Loc(n+1)$ = Location of consequent pixel, x = Multiplication Factor, $Loc(n)$ = Location of Previous Pixel Selected and Y = Incremental Factor M = Modulas Factor
4. Bit converter converts the selected pixel value into binary sequence $Bin(P)$ as per Eq. 22.2:

$$Bin(p) = \{b_{ij} | 0 < i <= r, 0 < j <= c \; and \; b_{ij} \epsilon (0, 1)\} \tag{22.2}$$

5. Bit extractor extracts the seven bits of $Bin(p)$ and form a sequence of bits as per Eq. 22.3 given below:

$$Bit(S, 7) = \{S(i) | 0 < i <= 7 \; and \; S(i) \epsilon (0, 1)\} \tag{22.3}$$

6. The bits sequence got in step 4 are added in augmented manner as per Eq. 22.4.

$$Aug(S) = \{S(i) + S(j) | i = 1, 1 < j <= 7 \; and \; 0 <= Aug(S) <= 7\} \tag{22.4}$$

7. Calculate the main value using Eq. 22.5:

$$Main(S) = \{Aug(S)\%2 | 0 <= Aug(S) <= 7 \; and \; 0 <= Main(S) <= 1\} \tag{22.5}$$

8. Insert bit b got in step 2 in sequence of bits compiled is step 4. If $b = 0$ and $Main(S) = 0$, then no alteration in pixel intensity is needed. If $b = 0$ and $Main(S) = 1$, then add or subtract 1 to the pixel selected such that $Main(S)$ becomes 0. If $b = 1$ and $Main(S) = 1$, then no alteration in pixel intensity is needed. If $b = 1$ and $Main(S) = 0$, then add or subtract 1 to the pixel value so that $Main(S)$ becomes 1.
9. Repeat all the steps until all the bits are inserted in the image. Generate the stego image $S(r, c)$ and exit.

22.3.2 Retrieval Algorithm

The seven steps of insertion algorithm are repeated at the receiver side on stego image $S(r, c)$. If $Main(S) = 0$ at the selected pixel value, then message bit is 0; otherwise, message bit is 1. The complete message will be retrieved by sequencing the all bits and translate them into alphanumeric characters.

22.4 Results and Analysis

The imperceptibility of the technique is tested by using practical checking by persons from different age groups. We have sent the cover image and stego image to many persons. We asked them about the difference between cover file and file image into five numeric values as per the following detail:

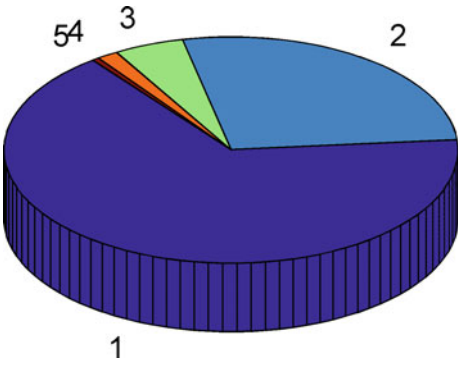


Fig. 22.3 Response submitted by different persons

1: No divergence amid cover file and stego file
2: Very minute divergence amid cover file and stego file
3: Average divergence amid cover file and stego file
4: Slightly high divergence amid cover file and stego file
5: Very high divergence amid cover file and stego file

We have got the response of 200 persons by using above numerical values on four different test images. The response submitted by different persons shown by pie graph on one test image in Fig. 22.3.

Mean value of responses is also calculated by using the following equation:

$$\text{Mean Value} = \sum_{k=0}^{200} \text{response}(k) \div 200 \tag{22.6}$$

Imperceptibility level is also classified into different levels as per different mean values which is given by Table 22.1.

The result of mean value on four different test images after insertion of 1 KB message is given below in Table 22.2. We have analysed the results from Table 22.2 and found that mean value lies between 1 and 2. So, we can say that the proposed method is highly imperceptible. We have also done the PSNR analysis on the above images for different size of messages. The calculation of PSNR is done by using the following equation:

$$\text{PSNR} = 10 \log_{10}(I^2 * \text{MSE}) \tag{22.7}$$

PSNR denotes peak signal-to-noise ratio, I denotes peak value of pixel, i.e. 255 in this case, and MSE is the mean square error given by Eq. 22.8. PSNR denotes peak signal-to-noise ratio, I denotes peak value of pixel, i.e. 255 in this case, and MSE is the mean square error given by the following equation:

$$\text{MSE} = \sum_{i=1}^{r} \sum_{j=1}^{c} (x_{ij} - y_{ij})^2 \tag{22.8}$$

Table 22.1 Imperceptibility level as per mean value

Mean value	Imperceptibility level
$1 \leqq$ Mean Value < 2	High imperceptibility
$2 \leqq$ Mean Value < 3	Average imperceptibility
$3 \leqq$ Mean Value < 4	Low imperceptibility
$4 \leqq$ Mean Value < 5	Very low imperceptibility

Table 22.2 Mean value of different test images

Image	Image size in pixels	Message size (KB)	Mean value
Building image	50,625	1	1.435
Cameraman image	16,384	1	1.625
Monalisa image	50,625	1	1.475
Ship image	50,625	1	1.730

PSNR on different images for different techniques is given by Table 22.3. The result shows that proposed method produces finer PSNR value than some state-of-the-art techniques. It provides acceptable value of PSNR. The comparison of PSNR values of different images for various size of message is also shown by Figs. 22.4 and 22.5. After analysing various graphs, we have found that proposed method provides better value of PSNR than some traditional methods.

22.5 Conclusion and Future Scope

The proposed method shows improvement over some state-of-the-art methods like LSB method, parity checker method, and GLM method. It provides improved value of PSNR. The subjective test also shows that the proposed method provides greater imperceptibility. Only one bit can be inserted in one pixel value of cover image. Our method also does not depend on LSB for insertion and extraction of the message. So, all the anomalies that are associated with LSB involved approaches are resolved by proposed method. LSB can be changed easily on the network by network noise. Intruder can also alter the LSB for destroying the message. All these types of problems are resolved by our method. In the future, we will try to develop some technique by increasing embedding capacity while retaining robustness and imperceptibility of the cover file.

Table 22.3 Mean value of different test images

Image	Message size (KB)	LSB method	GLM method	Parity checker method	Proposed method
Building image	1	35.51	37.54	38.30	45.20
Cameraman image	1	35.65	37.53	40.30	46.45
Monalisa image	1	35.89	37.02	41.39	46.46
Ship image	1	35.56	36.30	42.87	47.54
Building image	2	34.62	37.90	40.40	45.30
Cameraman image	2	34.65	37.45	40.98	45.29
Monalisa image	2	33.77	35.50	41.03	44.90
Ship image	2	33.31	34.20	39.07	45.67
Building image	3	32.10	36.30	38.70	45.09
Cameraman image	3	32.15	37.92	40.54	44.99
Monalisa image	3	33.17	37.80	40.34	44.78
Ship image	3	34.16	36.40	39.78	44.34

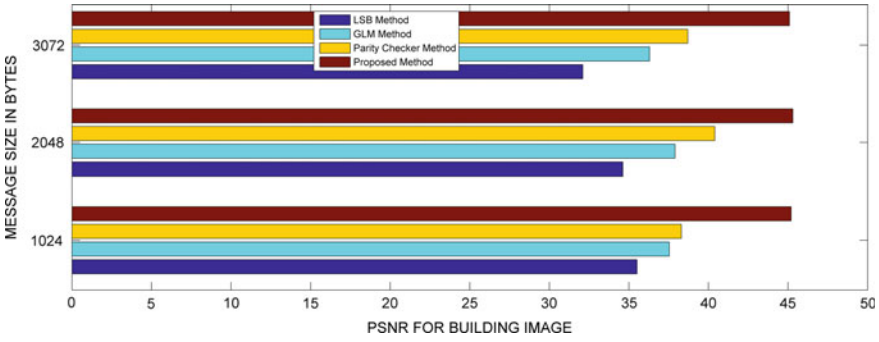


Fig. 22.4 PSNR value comparison of different methods for building image

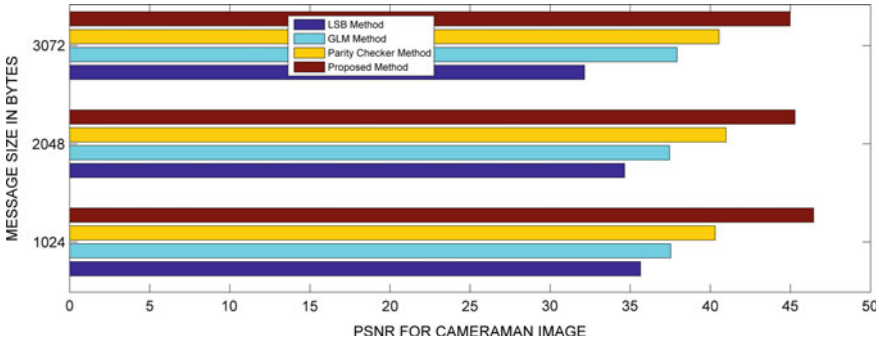


Fig. 22.5 PSNR value comparison of different methods for cameraman image

References

1. Gutub, A., Fattani, M.: A novel Arabic text steganography method using letter points and extensions. World Acad. of Sci. Eng. Technol. **1**, 502–505 (2007)
2. Krenn, R.: Steganography and Steganalysis. Internet Publication. Available at: http://www.krenn.nl/univ/cry/steg/article (2004)
3. Brainos, II, A.C.: A Study of Steganography and The Art of Hiding Information. Internet Publication. Available at www.infosecwriters.com (2007)
4. Provos, N., Honeyman, P.: Hide and Seek: introduction to steganography. IEEE **1**, 32–44 (2003)
5. Ahn, L.V., Hopper, N.J.: Public-Key Steganography. Lecture Notes in Computer Science. Springer, Berlin (2005)
6. Johnson, N.F., Jajodia, S.: Exploring steganography: seeing the unseen. IEEE Comput. **31**, 26–34 (1998)
7. Yadav, R., Rishi, R., Batra, S.: A new steganography method for gray level images using parity checker. Int. J. Comput. Appl. **11**, 0975–8887 (2010)
8. Potdar, V., Chang, E.: Gray Level Modification Steganography for Secret Communication. IEEE, New York (2004)

Chapter 23
CNN Architectures to Recognize Handwritten Telugu Characters



K. Dhana Sree Devi and C. Shoba Bindu

Abstract Artificial neural nets are backbones of many AIs predictive analogies. As common models in many of advancing domains: machine learning, artificial intelligence, and deep learning, they have led their own cognitive ways in developing highly expressive variants that work on image recognitions. From text mining to natural language processing, many of these domains have found neural networks in producing the expected accuracy. This work focuses on recognizing handwritten Telugu character vowels using convolution neural nets. We experimentally demonstrated the performance of CNN on varied architectures in predicting the handwritten Telugu vowels.

23.1 Introduction

The revival of neural networks for pattern recognition has made them to be widely used in image processing, today's most cognitive domain. Before neural networks, pattern recognition problems were most often addressed by KNN classifiers [1], quadratic discriminates [2]. Later with the advent of neural networks, image classification is more facilitated with many of its sophisticated variants. Neural networks with their activation functions [3] and back-propagation algorithm are the beacons in artificial intelligence and deep learning [4] and are almost successful in delivering new insights in image processing. Fourier transformation and linear discrimination approaches are early choices of image recognition and reconstruction. Despite their strong mathematical theories, few weaknesses still persist in LDA [5]. Many other versions of discriminate analysis are developed in pattern recognition [6].

State of the art on neural network predicted them as underlying image recognizers and their prediction was so accurate, that the field started using them in solving many

K. Dhana Sree Devi (✉)
Department of CSE, CVR College of Engineering, Hyderabad, Telangana, India

C. Shoba Bindu
Department of Computer Science Engineering, JNTUA College of Engineering, Anantapur, Andhra Pradesh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_23

real-time problems. Their success in image processing was truly because of their well-defined architecture and well-trained on the taxonomy of image processing. Neural network back-propagation is the key contributor to the image compression technique [7]. Their ability to reconstruct the original image [8] in linear time is also a point of conquer. With the advent of machine learning and deep learning, the predictive usage of neural nets grew in a way that led to the development of many of its variants for image classifications. The convolution neural nets, recurrent neural nets, LSTMS, and GANs are popular variants [9–11]. All of these variants have in common neural network architecture of I/O, hidden layers wherein their approaches and purposes are different.

Deep neural networks (DNNs) difference themselves with more hidden layers. Convolution neural networks (CNNs) are a kind of DNNs that are feed-forward [12]. Ordinary neural network and CNN mend a difference in image classification. CNNs are widely used in image processing because of their adaptability and usage of few training parameters. Image classification renders another special kind of neural networks the recurrent neural nets, which showcase a special sequencing operation on inputs. These sequences are used to learn the patterns over time. RNNs architecture uses a connection from output to input which enables to gain a temporary memory [13] of the past data. RNNs are successful in keeping the track of long-term dependencies of the Input sequences. The core of RNN is the back-propagation algorithm due to which the drawback of vanishing gradients do persist in RNNs.

Long short-term memory (LSTM) is improved version of RNNs that have more prominent usage in deep learning. LSTMs were developed to deal with the RNNs drawback of exploding and vanishing gradients [14]. More often LSTMs are used in image labeling, speech recognition by many of major technology companies Apple, Google, and Microsoft.

23.2 Architectural Study of CNN

In this section, we are going to present the architectural components of CNN and their major computational units rendering to performance raise in image recognition.

23.2.1 General Artificial Neural Network

General neural networks which are the backbones for the development of major variants are networks with various layers: the input, output, and a middle hidden layer. The layers include the computational units called the neurons, and these neurons of one layer extend a connection to other layered neurons. Each neuron acts as a multiplier: multiplying the incoming input with the network weights and sums these.

The architectural key aspects are:

- Role of activation function
- Role of bias
- The learning process
- Parameters to improve the image recognition performance.

23.2.1.1 Role of Activation Function

The type of activation function used defines the predictive power of any NN. Simulation using nonlinear activation is practiced by many complex neural networks. To identify the erroneous behavior of the data, the choice is always nonlinear activation functions [15] which can adapt to the nonlinear characteristics of the inputs. Some of the activation functions used are:

- Sigmoid
- Tanh
- ReLU
- Softmax.

Linear functions always do exhibit a constant slope. Slope as we know pictures the rate of change of the predictor variable (y) with respect to the independent variable (x). Real-world problems emphasize on finding the points of minimal or maximal change: may be minimization of loss or maximization of gain. If we take linear activation functions, these changes are constant and cannot be distinguishable between points of minimal and maximal changes. Nonlinear activation functions have gradients that vary between various points. Based on these gradients descent or accent, we can address all the minimization and maximization problems. The nonlinear exponential function e^X has a range between [0, ∞] and has an infinite range. The gradients for the function e^X exist between [0, ∞]. To find the optimal gradients, it is quite difficult within this infinite range. Hence, such functions though nonlinear cannot be taken as activation functions. Hence, the needs for activation functions which show active gradients within shorter intervals are needed. One such is the sigmoid activation function which will analyze the gradients between a small [0, 1] interval.

23.2.1.2 Role of Bias

Bias is one more important parameter in an NN. The output from a neuron is of the form: O/P = $\sum$(weights*Inputs) + Bias.

This is more similar to a linear function: $Y = mX + C$; here a positive or a negative constant C shifts Y right or left; even the bias does the same. Weights of the NN are used to steer the steepness of activation function; bias will shift this steepness to the right or left of the curve [16]. These shifts are indications of the delayed triggering of the activation function. Thus, bias is used to preserve the trigging active nature of the activation cell even for zero inputs. If the NN is trained under zero bias, it does not

mean that bias cannot be zero. There are NNs which work for zero bias too. But in reality all NNs use bias. Bias is critical for successful learning of the networks; with them the networks can learn different outputs for inputs. A look through Fig. 23.1 reflecting triggering of sigmoid with only weights and without bias. The plot shows steepness in the curves as weights are varying. This is a clear indication that the model is learning fast under zero bias. Neural network fast learning [17] may surpass some of the converging points.

A close look through Fig. 23.2 where a bias term is varied shows the slow triggering of the sigmoid. The moderate learning by the model may analyze various true converging points. Hence, bias terms play a critical role in analyzing the converging points of the model.

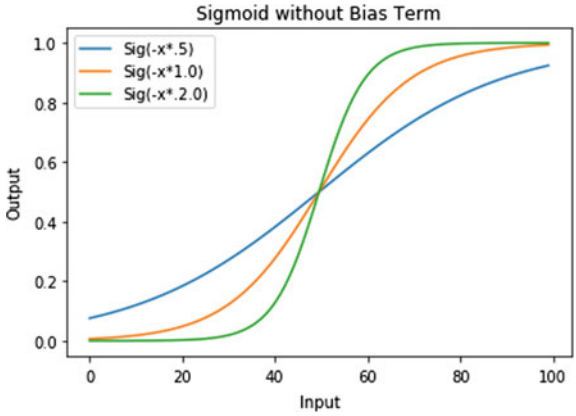


Fig. 23.1 NN learning under zero bias

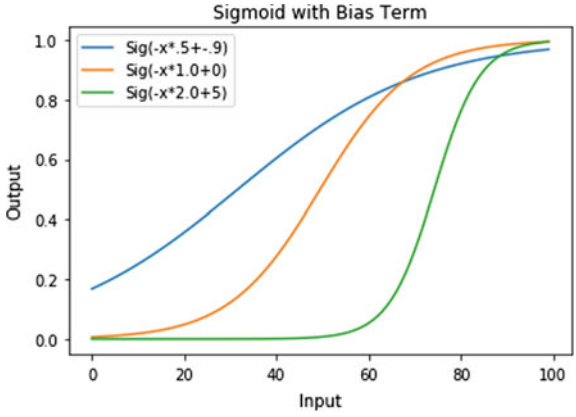


Fig. 23.2 NN learning under some bias

23.2.2 Convolution Neural Networks

They are completely different from other NN variants. Falling into the class of deep neural nets with more hidden layers is most commonly used for image classification. They are different by using high-layered structure with convolution layers, pooling layers, and the classification layer. In usual neural networks, few hand engineered filters are used to accomplish the preprocessing stage; ConvNet need less preprocessing as they have the ability to learn these filter characteristics. The first practically used ConvNet is LeNet-5 [18, 19]. Figure 23.3 shows the architecture of CNN.

The architectural key divisions are:

1. Feature extraction in multiple hidden layers
2. Classification in the output layer.

Feature extraction from CNN involves the following operations:

- Convolution
- Nonlinearity (ReLU)
- Pooling or Sub-sampling.

Whereas classification in the output layer includes:

- Classification using Fully Connected Layer.

23.2.2.1 Feature Extraction in CNN

Before feature extraction little bit of preprocessing is done by normalizing and sizing of image data. Image feature extraction is a kind of dimensionality reduction. Generally, an image is represented as pixels each containing 8-bit information. If the image is too large, this pixel information size is also large and is too much for a model to be trained with. So they initially undergo a feature extraction stage where distinct key parts of the image like corners, edges, and intense patches are identified.

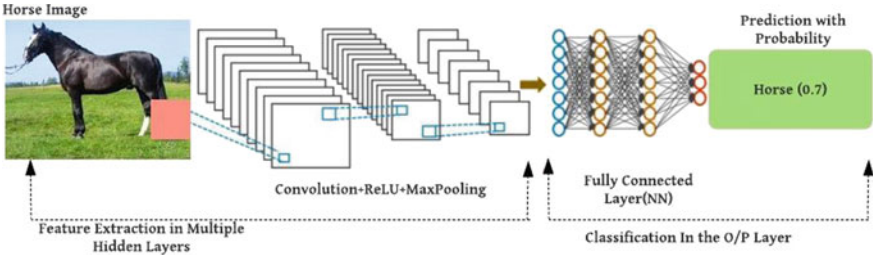


Fig. 23.3 ConvNet architecture

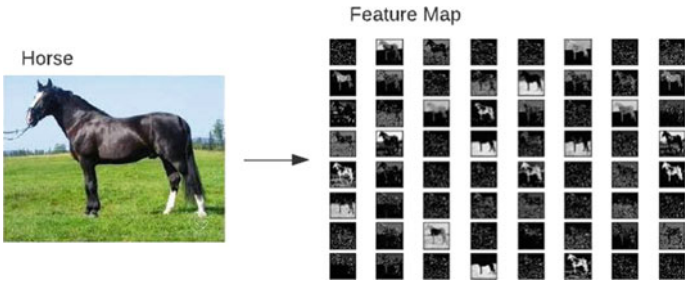


Fig. 23.4 Image feature map

Convolution

The convolution operation anneals in passing the preprocessed data through a convolution layer [20]. Here input image is condensed by extracting key features of interest so as to produce various feature maps. The convolution layer uses various filters in sequence to extract the stack of feature maps. Each filter is convolved with the image to extract the sequence of low level, middle level, and finally the high level feature maps. Figure 23.4 shows feature map extracted from an image.

Each convolution layer is stacked with various parameters; I/0 sizes, depth of feature map, kernel size, and stride. A filter which is a vector of learnable weights is used to convolve the input. These weights are learned using a back-propagation algorithm. Later these learned filters are applied to the input images to create feature map; feature map which summarizes the presence of those features in the image input. Figure 23.5 shows 5×5 filter applied to convolute an image of size 7×7.

Figure 23.6 shows 5×5 image convoluted with 3×3 filter to obtain 3×3 vector feature map. Here the size of the vector image is 5×5 and the size of the filter is 3×3, as the image vector is larger than the filter vector the filter is slide on the image vector as shown in Fig. 23.5. This generates a new output pixel. The slide length, which is the number of pixels we slide, is called as the stride. Figure 23.6 shows a one stride output vector where the pixel value 6 is obtained by convolving the filter with 3×3 image sub vector. Similarly by striding all other output pixels are obtained. The filter used in Fig. 23.6 is an identity filter. There are other filters too. Once the feature maps are obtained, nonlinearity transformation using ReLU is applied on these feature maps.

The output feature map dimensions after convolving is given by:

$$O = ((W - K + 2P)/S) + 1 \tag{23.1}$$

where

W is the width of the input feature map.

K is the width of the filter.

P is the padding layer dimension.

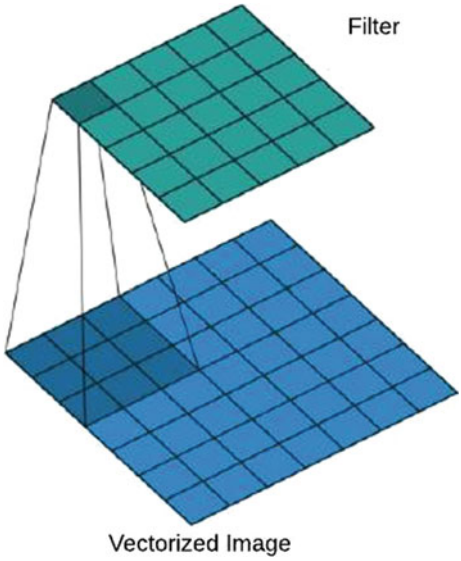


Fig. 23.5 5×5 filter to convolute 7×7 image

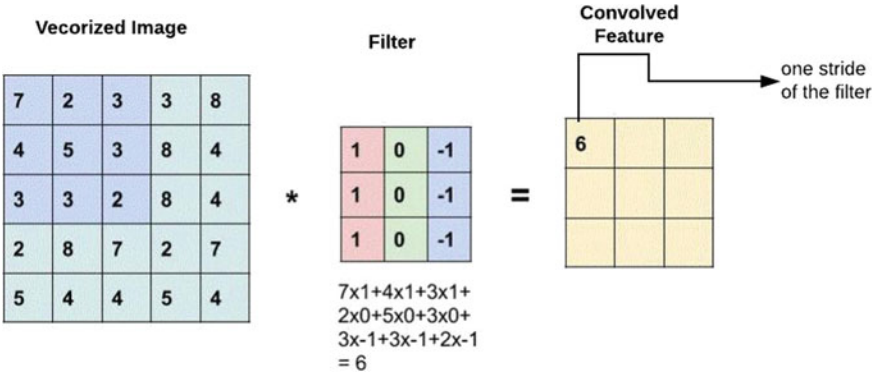


Fig. 23.6 Convolute to extract feature map

S is the stride.

Pooling Layers: The major drawback with the convolution layers is that they mark and record, the exact positions of the features in the input. An indication that small shifts in the feature position of the original input image results in a different feature map. Down sampling [21] is a common approach to address this problem and is achieved by adding pooling layers. These pooling layers are typically added after the convolution layers and after ReLU nonlinearity transformation. Pooling layers does not need a training back-propagation as their operations are fixed. Pooling involves

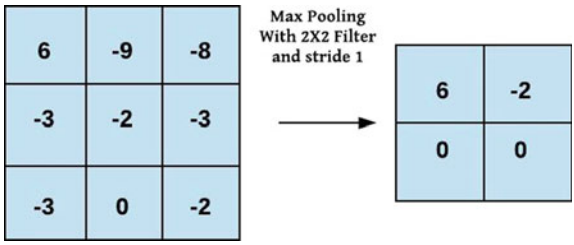


Fig. 23.7 Applying max pooling using 2 × 2 pooling filter

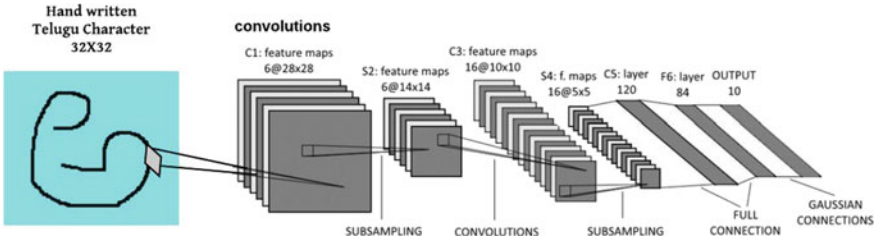


Fig. 23.8 CNN-LeNet-5 structure with Telugu vowel input

in choosing a suitable pooling operation, and two common pooling operations are given by:

1. Average pooling
2. Maximum polling

For each patch on the feature map, average pooling calculates the average value, whereas the max pooling calculates the maximum value. Thus using either average or maximum approaches produces a pooled feature map with the image feature constrained to the same location. This might be otherwise called as a dimensionality reduction approach by the pooling layers in feature maps of CNNs. The 3 × 3 output feature map in Fig. 23.8 is pooled using Max pooling approach with a 2 × 2 filter and using stride 1 and is shown in Fig. 23.7.

23.2.2.2 Classification in the Output Layer

Classification layer is normal fully connect NN network as discussed in Sect. 23.2.1. It collects the final convoluted feature from the convolution layer and returns a column vector of classes, and with class estimations to which the image belongs as shown in Fig. 23.3.

23.2.2.3 Parameters to Improve the Performance of Image Recognition in CNN

When it comes to deep learning models, performance is a key aspect to be considered. Model accuracy is an important parameter to be brooded over. Once the model is built and when its performance is not meeting the desired, then tuning the performance comes into picture. What we can do to improve the performance of a model is to use various performance improvement techniques. Accuracy of CNNs can be improved in several ways:

- Using a hyper-parameter tuning
- Using improved data.
- Use regularization techniques.

Improving Accuracy with Hyper-Parameter Tuning

In general, the hyper-parameters are specific to the model what we have built and that we want to optimize. These can also be discussed as the settings that are designed to control the model behavior. Some classic hyper-parameters to think of in boosting the performance are:

Number of Hidden Neurons: Number of hidden neurons affects the model performance. With few hidden neurons, the model may fail to learn underlying patterns of the image. So as a first aspect to boost the performance we change: either decrease or increase the hidden neuron count. But this may sometime outperform in test data and may cause to under-fitting or over-fitting, a model with different biased errors. We have to tune this parameter to have an appropriate capacity of fitting the data. Accurate predictions do not always tag a good model; but the one which gives accurate predictions on new data is needed.

Cross-validation: A method is used to estimate the model performance on new data. Model improvement prefers a threefold or a fivefold cross-validation where the entire data is divided into 3 or 5 sub-parts in counter to 70–30 train–test split on the whole data. For example in a threefold cross-validation, train data is divided into 3 subsets T1, T2, T3. In the first run, we fit the model using T1, T2 and use T3 for validation and so as for the next runs and the final accuracy is the mean of the accuracies in each run. When k-fold cross-validation is implemented at one point each and every subset of k-fold is considered new validation data on which the performance is measured. Many works, experimentally showed k-fold cross-validation produced less biased estimate. Here the hyper-parameter is k the number of folds, which can be chosen manually or can be tuned using various parameter tuning approaches, which we will discuss in the later sections.

Using Improved Data

A great challenge is prepare the best data to train the CNN. When we are working on image data set, truly all the images may not be of same kind and size. As to raise the performance, a common approach is preprocessing the data. Various image data preprocessing approaches are: Progressive resizing, Batch Normalization. More often data can be very small then Image augmentation can be used to increase the training data size and can also be discussed as one of the regularization approach to address over-fitting.

Progressive resizing is one of the performance boosters where the images are iteratively resized to analyze the performance. Initially the model is trained with a smaller image size may be 64×64. A transfer learning is then used where the weights obtained in the first training phase with 64×64 image are used to train the model on image sizes 128×128 under the same layered architecture.

Normalization is often done before the model is built in many of the cases; in CNNs normalization takes place inside the network. Under normalization, all the image pixel values are made to fall between 0 and 1 or other ranges. Normalization can scale the data to smaller ranges so that it is easy to tune an optimal learning rate.

One among widely used techniques to improve the CNN performance is Image augmentation. As one of a kind of data augmentation approach, image augmentation generates more of new training samples from the existing and the CNN can be trained with more of newer data. The approach proved to be very useful when we have small training data. Augmentation is where small sub-sampled images are considered from the original image. More advanced Image augmentations like Copyout, CopyPairing showed an improvement in test accuracy. The size of sub-sampled image is one of the hyper-parameter which can be tuned.

Using Regularization Techniques

Most common approach to boost the CNN performance is to avoid over-fitting problem. Regularization is an approach for performance tuning where the learning algorithm is modified so as to generalize better. CNNs use two popular regularization techniques: dropout and early stopping.

As CNN with neuron nodes are concerned, dropout is the most efficient regularization approach, producing optimal results. As known a fully connected NN is where every neuron is connected to every other layered neuron; for a deep network, the structure may be even too large. Also all the neurons of the network may not contribute in meeting the loss function. Such neurons may drop performance. Dropout is where such neurons are identified using a threshold, and these neurons are dropped from the network. On dropping these neurons, the neuron incoming and outgoing connections are also removed and the network becomes smaller. So dropout has become a general choice of regularization for deep networks. The number of neurons to be dropped can be taken as a hyper-parameter.

23.3 Handwritten Telugu Character Recognition

Our work presents the simulation of classical CNN, the LeNet-5, and its two varied architectures, in recognizing the handwritten Telugu vowels and demonstration of experimental performance comparison by tuning various parameters. To add understandings on CNN layered parameters, this section discusses the classical LeNet-5 architecture in detail.

23.3.1 Classical LeNet-5 Architecture

Figure 23.8 shows a classical LeNet-5 CNN learning architecture and is a widely used CNN architecture for handwritten character recognitions. Trained with back-propagation algorithm, the network led for the advent of many advanced CNN architectures. LeNet-5 architecture includes a group of 7 layers, other than the input layer. Three convolution layers C1, C3, C5, two sub-sampling layers S2, S4 and one fully connected layer, one output layer. A 32 × 32 normalized input of a Telugu character image is provided to the network. The images are resized to guarantee better recognition of the features at the center.

23.3.2 Proposed ConvNet Architecture

A follow up of 3 key changes is done in our proposed architectures, shown in Fig. 23.9:

1. Changing the LeNET layers
2. Using ReLU and ELU activations
3. Improving the model training with improved dataset.

We experimented on two datasets: the normal data set and the improved dataset. As Telugu language vowels are more different from other language characters and

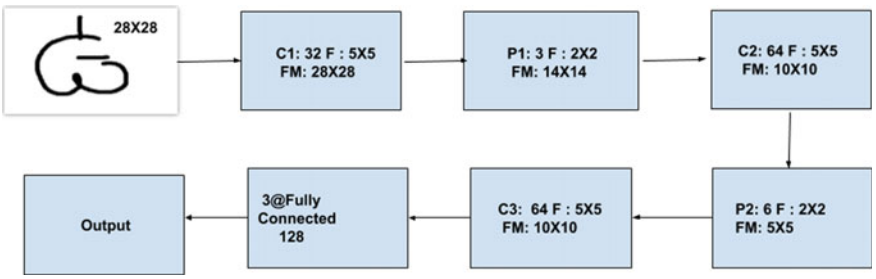


Fig. 23.9 Proposed model architecture

we used two models where we changed the architectural layers and the activation functions. Our model 1 comprised of 3 convolution layers using ReLU activation and a kernel size of 5 × 5. After the last convolution layer, we used a maximum pooling layer with size 3 × 3, the output of which is fed to fully connected layers for image flattening. The final layer is the output layer with 15 nodes and using Softmax and will classify 15 labels. Model 1 is build using normal data what we have initially.

Our model 2 architecture is build to reduce the misclassification error. We changed the model architecture with different layers and output nodes. Here we took 4 convolution layers using ELU activation and with 3 × 3 kernel. We took the same max pooling of 3X3. When experimenting on model 2, we took improved dataset.

We have chosen dropout as regularization approach, a varying dropout from 20 to 30% is what we want to experiment with for both the models.

23.3.3 Data Preparation

Telugu characters are huge in number an actual count of 56 and only 52 of them are more frequently used in language building. Our work included handwritten character classification of 15 vowels. A sample of which is shown in Fig. 23.10. We used English labeling for these 15 vowels accordingly on how they are pronounced. For each character, more than 500 handwritten images are created each of size 200 × 200, to include our normal dataset size of 7000 images.

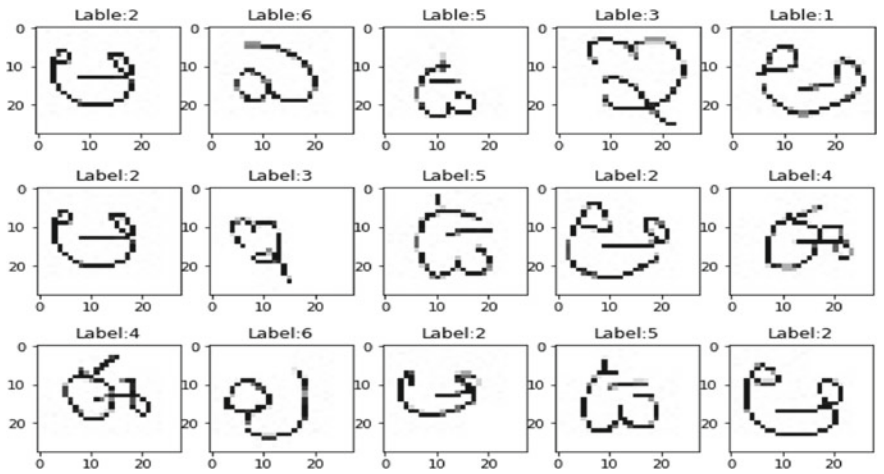


Fig. 23.10 Training samples

23.3.3.1 Improved Data Set

We used data augmentation to raise our sample size. Augmentation added new images to the base dataset using various transformations: image rotation, color change, adding noise. We used small rotation on about 5° as Telugu characters are scripted always with small inclinations from the zero line. Even the image shifts what we used are small of about 5% both vertical and horizontal. We applied ZCA whitening technique.

23.4 Experimental Results

The classical LeNET CNN is the reference for many experimental works, giving a scope to many research ideas. As a part of our proposed work, we modeled Telugu vowel classification on these three models: the LeNET and our proposed architectures model 1 and model 2. We first trained the LeNET to classify the vowels. We trained our model with small data first; the data what we have initially, 7000 images. We observed the accuracy rate at various dropouts; but the misclassification error is too large with 40%.

We then took our proposed model 1, trained with the same 7000 images. We used ReLU activation here. Here also we observed the misclassification error is too large with 40%. Here we observed that two of the vowel classes 'Ee' and 'A' are totally misclassified; as our trained data in these two classes is very less. We switched to our proposed model 2 using ELU activation, and with improved dataset. Using ELU decreased misclassification rate to some extent 34.33%. Here we observed that some of the vowels from 'Ee' and 'A' are correctly classified. We observed that improved dataset with augmentation performed on the data has reduced the misclassification error. We observed the misclassification error on varied dropouts.

In the second stage of experiment, we used our larger dataset which is improved using data augmentation as discussed in Sect. 23.3.3.1. We ran all the three models again by varying the dropout. Model 2 which used ELU activation performed well at dropout 25%; the misclassification error of 30.62%, which is less compared to other dropout observations on other models. Table 23.1 shows the misclassification rates and performance comparison of our two models with classical LeNET. Figure 23.11 shows some misclassified images.

23.5 Conclusions

Native language processing is emerging with new techniques and models with which the linguistic characters are being trained for further recognition. Among such models, the convolution networks are predominantly in use showcasing their accuracy in correctly recognizing the handwritten characters. Convolution neural

Table 23.1 Misclassification error %

Model	Test misclassification error % (Normal dataset)			Test misclassification error % (Improved dataset)		
	Dropout 20%	Dropout 25%	Dropout 30%	Dropout 20%	Dropout 25%	Dropout 30%
LeNET	40.80	38.72	42.78	37.68	35.21	39.55
Proposed model-1	40.01	39.41	43.11	36.91	35.00	38.97
Proposed model-2	37.33	34.76	38.24	33.33	30.62	36.60

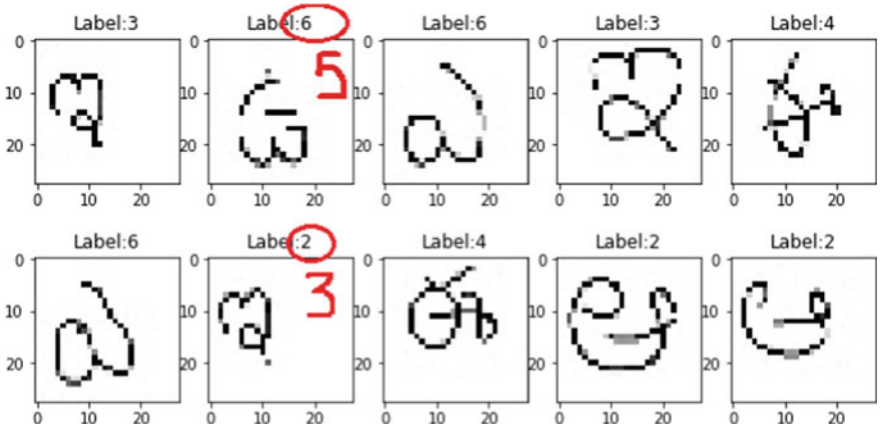


Fig. 23.11 Some misclassified images

networks are deep learning models trained with more of image data. With convolution filters, these models are able to extract recognizable features in images with more predictive powers. Natural language processing is much benefitted using this model. Our work showed one such case of experimental study of pre-trained CNNs in recognizing handwritten characters from Telugu language.

References

1. Vajda, S., Santosh, K.C.: A fast k-Nearest Neighbor classifier using unsupervised clustering. In: International Conference on Recent Trends in Image Processing and Pattern Recognition (Apr. 2017)
2. Bose, S., Pal, A., SahaRay, R., Nayak, J.: Generalized quadratic discriminant analysis. Pattern Recogn. **48**(8), (2015)
3. Dhana Sree, K.: Data analytics: role of activation function in Neural Net. IJITEE **8**(5), (2019)
4. Ongsulee, P.: Artificial intelligence, machine learning and deep learning. In: International Conference on ICT and Knowledge Engineering. IEEE (Nov. 2017)

5. Liang, Y., Gong, W., Pan, Y., Li, W.: Generalizing relevance weighted LDA. Pattern Recogn. **38**(11), (2005)
6. Wen, Y., Shi, Y.: An approach to numeral recognition based on improved LDA and Bhattacharyya distance. In: International Symposium on Communications, Control and Signal Processing. IEEE (2008)
7. Jiang, J.: Image compression with neural networks-a survey. Signal Process.: Image Commun. **14**(9), (1999)
8. Zhou, Y.T., Chellappa, R., Vaid, A., Jenkins, B.K.: Image restoration using neural network. IEEE Trans. Acoust. Speech Signal Process. **36**(7), (1988)
9. Angeline, P.J., Saunders, G.M., Pollack, J.B.: An evolutionary algorithm that constructs recurrent neural networks. IEEE Trans. Neural Netw. **5**(1), (1994)
10. Gers, F.A., Schmidhuber, J., Cummins, F.: Learning to forget: continual prediction with LSTM. ICANN, (1999)
11. Zhang, Y., Zhou, Y., Huang, C.: Generating artificial images by generative adversary network. In: International Conference on Pattern Recognition and Artificial Intelligence (2019)
12. Nebauer, C.: Evaluation of convolutional neural networks for visual recognition. IEEE Trans. Neural Netw. **9**(4), (1998)
13. Sukhbaatar, S., Weston, J., Fergus, R.: End to end memory networks. Adv. Neural Inf. Process. Syst. **28**, (2015)
14. Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks. J. Mach. Learn. Res. (2013)
15. Li, H., Ouyang, W., Wang, X.: Multi bias nonlinear activation in deep neural networks. J. Mach. Learn. (2016)
16. Tsujitani, M., Koshimizu, T.: Neural discriminant analysis. IEEE Trans. Neural Netw. (2000)
17. Fu, L.M., Hsu, H.H., Principe, J.C.: Incremental backpropagation learning networks. IEEE Trans. Neural Netw. **7**(3), (1996)
18. Han, S., Pool, J., Tran, J., William D.: Learning both weights and connections for efficient neural network. Neural Inf. Process. Syst. (NIPS) (2015)
19. AI-Jaw, R.: Handwriting Arabic character recognition LeNet using neural network. Int. Arab J. Inf. Technol. **6**(3), (2009)
20. Jeon, Y., Kim, J.: Active convolution: learning the shape of convolution for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition (2017)
21. Sun, M., Song, Z., Jiang, X., Pan, J., Pang, Y.: Learning pooling for convolutional neural network. Neurocomputing **22**

Chapter 24
Rider-Deep Belief Network-Based MapReduce Framework for Big Data Classification



Sridhar Gujjeti and Suresh Pabboju

Abstract Big data is an emerging domain with huge amounts of data and is increasing rapidly in an exponential manner. There exist different eras in the massive datasets, but the analysis of huge data in the medical field is a challenging task. This paper presents a Rider-based Deep belief network (Rider-DBN) for classifying big data in the medical domain to predict diseases. The goal is to provide a MapReduce framework with a deep learning classifier that can categorize normal and abnormal classes with improved accuracy. Here, the information gain is employed for selecting the imperative features using inputted data considering mappers. Thus, the dimension of the features is reduced in such a way that only the selected features are given for the classification to handle the big data. The selected features are employed by the Deep belief network (DBN) for categorizing the normal or abnormal classes in disease prediction. The Rider optimization algorithm (ROA) is employed to train DBN that finds optimum weights of the DBN for disease prediction. The Rider-DBN provides superior performance with the accuracy of 89.3%, sensitivity of 78.6%, and specificity of 92%.

24.1 Introduction

The evolution attained in computer technologies, digital data storage, and scientific encroachment in the communication domain has facilitated the cohort of massive datasets in medical applications. The extraction of patterns using medical data assists the clinicians to diagnose the patients [1]. Many of the hospitals utilize the information of hospital patients and their histories in a storage repository for diagnosis, but these methods may maximize the size of storage and degrade the performance of the system [2]. Healthcare is suffering from several issues for detecting the patient's

S. Gujjeti (✉)
Kakatiya Institute of Technology and Science, Warangal, Telangana, India

S. Pabboju
Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_24

syndrome [3]. The classification of disease using the attributes is a noteworthy research area from a diagnosis standpoint. The effective classification is attained using a suitable mechanism which capitulates improved accuracy [4]. In [5], a rough-fuzzy hybrid classifier is devised for predicting cardiac disease. In [6], SVM was utilized for categorizing tumors based on the Factorizing of Symmetry-Non Negative Matrix. In [7], a genetic-fuzzy system is devised using ECG signals for heart disease diagnosis. In [8], a model is devised for diagnosing cardiac ailment using an optimization algorithm. A neural network [9] is devised for classifying the human brain tumor using magnetic resonance images. In [10], the method was devised for monitoring type 2 diabetes mellitus using patient information.

The major contribution of the paper:

- **Proposed Rider-DBN for disease prediction using massive-sized data**: The classification of big data is carried out using DBN for predicting normal and abnormal classes using input disease data. The training of DBN is done by ROA, which tunes the optimal weights for discovering the classes.

24.2 Literature Review

The four classical disease prediction strategies using big data are illustrated along with their merits and demerits. Lakshmanaprabu et al. [11] devised an IoT-based healthcare system based on Random Forest Classifier (RFC) with MapReduce for big data classification to predict the disease. The method showed effective performance, but the algorithm was computationally slow due to massive datasets. ALzubi et al. [12] devised Weight Optimized Neural Network with Maximum Likelihood Boosting (WONN-MLB) by analyzing huge data for disease prediction. The method showed high diagnosis accuracy, but the method did not apply to other datasets. Nalluri et al. [13] devised a hybrid model by combining SVM and monarch butterfly optimization (MBO) for disease prediction. This method attained effective tradeoffs amongst cataloging three goals, which involve specificity, accuracy, and sensitivity. Meanwhile, the technique was not capable to incorporate multi-objective bi-clustering with parameter evolution. Reddy et al. [14] devised an adaptive genetic algorithm with fuzzy logic (AGAFL) for predicting heart disease and helping the clinicians to spot heart disease in an earlier stage. The method was effective in handling noisy data and poses the potential to deal with a large number of attributes.

24.3 Proposed Rider-DBN for Predicting Disease Using Big Data

Figure 24.1 portrays the architecture of the proposed Rider-DBN for disease prediction using big data. Here, the input data is fed to the MapReduce framework wherein

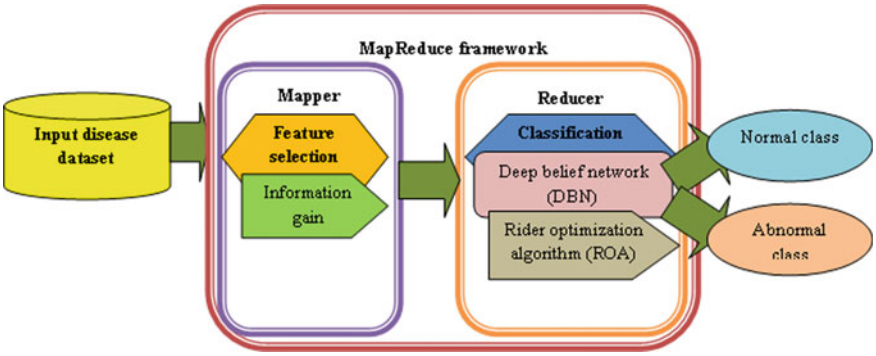


Fig. 24.1 Schematic view of proposed Rider-DBN for predicting disease in big data

the selection of imperative features and classification is performed to predict disease. The feature selection is performed in the mapper function using Information gain and the classification is performed on reducer with DBN, where the training is performed using the ROA that determines the weights of the DBN. Thus, the DBN classifier in the reducer performs the disease prediction and generates two classes, normal or abnormal classes.

24.3.1 Selection of Imperative Feature Using Information Gain in Mapper Phase

The map-reduce model has elevated power for processing since different servers are processed in parallel and poses the capability to perform parallel processing. Consider the input big data is expressed as D with various attributes as given in Eq. (24.1),

$$D = \{d_{e,f}\};\ (1 \leq e \leq P);\ (1 \leq f \leq S) \tag{24.1}$$

where $d_{e,f}$ signifies data in big data D with eth data of fth attribute, P denotes total data points, and S represents total attributes. The mapper adapts the mapper function, and ROA is used to select the best features. The importance of selecting features is that imperative features are selected for ensuring dimensionality reduction and to elevate the classification accuracy. Initially, big data is splitted into various subsets, which is represented in Eq. (24.2).

$$d_{e,f} = \{t_u\};\ (1 \leq u \leq F) \tag{24.2}$$

where, F denotes total subsets. Besides, total subsets are equivalent to total mappers and expressed in Eq. (24.3).

$$M = \{M_1,\ M_2, ...M_u, ..., M_F\} \tag{24.3}$$

The input to the uth mapper is represented in Eq. (24.4).

$$a_u = \left\{B_{i,\,j}\right\};\ (1 \leq i \leq K);\ (1 \leq j \leq S) \tag{24.4}$$

where, S represents total attributes and K denote total data points in uth subset such that $K < P$. The mapper employs information gain [15] for choosing optimum features and output from all mapper is combined to produce chosen features. The information gain is a measure employed for selecting optimum features. Information gain is also termed as Kullback–Leibler divergence which can be utilized as a measure to rank each feature. For computing information gain, firstly the entropy of the dataset must be computed. The entropy of the dataset D is computed using Eq. (24.5).

$$E(D) = -\sum_{m \in J} P(m) \log P(m) \tag{24.5}$$

where, $P(m)$ represent the probability of instance $d_{e,f}$ to be labeled as m and it is represented in Eq. (24.6).

$$P(m) = \frac{|\{d_{e,f} \in D | q_{e,f} = q\}|}{|D|} \tag{24.6}$$

where, $d_{e,f}$ denote the data in big data D representing eth attribute of the fth data, D is total subset of data, and $q_{e,f}$ indicate label of eth attribute of the fth data. The information gain of lth feature of the dataset D is computed using Eq. (24.7), where the first part indicates the sum of probabilities of the instance $d_{e,f}$ to have a value n of the feature. The second part represents entropy of a subset of instances from D that has the value n of lth feature and is formulated as,

$$\mathrm{IG}(D, l) = E(D) - \sum_{n \in N^l} \frac{|\{d_{e,f} \in D\} | d^l_{e,f} = n\}|}{|D|} E(\{d_{e,f} \in D | d^l_{e,f} = n\}) \tag{24.7}$$

After the application of information gain, the obtained selected features are represented as a feature vector A. The chosen features are subjected as input to reducer, which process to produce needed output with the proposed Rider-DBN model.

24.3.2 Classification of Massive Data with Proposed Rider-DBN in Reducer Phase

The proposed Rider-DBN is employed in reducers to achieve classification by offering optimal classification results to guarantee the effectual extraction of the

pattern using big data and to reduce the time taken for processing. Moreover, the reducer is expressed in Eq. (24.8).

$$V = \left\{V_1,\ V_2, ..., V_y, ...V_z\right\} \tag{24.8}$$

where, z represents total reducers where $(1 \leq y \leq z)$. The evaluation of classification results in reducer is adapted for performing effectual classification.

24.3.2.1 Training of DBN Using ROA Algorithm

The DBN [16] is a subset of Deep Network and comprises various layers of Restricted Boltzmann Machines (RBMs) and Multilayer Perceptrons (MLPs). The training of DBN is carried out using the ROA [17] that aims to discover the optimum weights to tune the DBN classifier. The classification of big data employs proposed Rider-DBN to classify data by obtaining the best classification and is capable to handle new data attributes. The steps of the ROA algorithm are depicted in this section.

Step (1) Initialization: At first, the riders are randomly initialized and expressed in Eq. (24.9).

$$Y = \{Y_k\}\ ;\ (1 \leq k \leq l) \tag{24.9}$$

where, l represents the total number of riders, and Y_k indicates kth rider from the total riders.

Step (2) Evaluate the error: The optimum solution is discovered using fitness function and is considered as a minimization function and hence, the solution with the least Mean Square Error (MSE) is selected as an optimal solution and MSE is formulated in Eq. (24.10).

$$MS_{err} = \frac{1}{P}\sum_{e=1}^{P}\left[O_e - O_e^*\right]^2 \tag{24.10}$$

where, O_e is expected output and O_e^* is the predicted output, P indicate total input data where $1 < e \leq P$.

Step (3) Update the solution based on the ROA algorithm: The solutions are updated after measuring the error and the solution update is performed with ROA using four-rider parameters, namely follower, attacker, overtaker, and bypass rider.

The follower updates position by the leading rider location for reaching the target in rapid mode and is formulated in Eq. (24.11).

$$Y_{s+1}^f(k,l) = Y^N(N,l) + [\text{Cos}(W_{k,l}^s * Y^N(N,l) * D_k^s)] \tag{24.11}$$

where, l is coordinate selector, Y^N signifies leading rider position, N denote leading rider index, $W^s_{k,l}$ represent steering angle of kth rider in lth coordinate, and D^s_k is distance.

The updated position of overtaker is employed to maximize the rate of success by discovering the overtaker position and expressed in Eq. (24.12).

$$Y^o_{s+1}(k,l) = Y_s(k,l) + \left[D^*_s(k) * Y^N(N,l)\right] \tag{24.12}$$

where, $D^*_s(k)$ represent the direction indicator.

The attacker grabs the leader's position by the following leader and the attacker update position is formulated in Eq. (24.13).

$$Y^a_{s+1}(k,b) = Y^N(N,l) + [\text{Cos}\, W^s_{k,b} * Y^N(N,l)] + D^s_k \tag{24.13}$$

The bypass riders undergo a common path without following the leading rider and are represented in Eq. (24.14).

$$Y_{s+1}(k,b) = \alpha[Y_s(\mu,b) \times \eta(b) + Y_s(\vartheta,b) * [1-\eta(b)]] \tag{24.14}$$

where, α denote a random number between [1, 0], μ is a random number in the interval [1, P], ϑ is a random number between [1, P], η is a random value, b is the coordinate, and k denote the position of the rider, and $Y_s(\mu, b)$ represents the position of μth rider in bth coordinate.

Step (4) Re-computation of error: The error is calculated using Eq. (24.10) and the solution with minimal error is used to train DBN to detect disease.

Step (5) Terminate: The optimum weights are obtained in repeated mode till utmost iterations are attained.

24.4 Results and Discussion

The assessment of the proposed strategy with classical strategies using sensitivity, accuracy, and specificity is illustrated. The dataset employed for experimentation is Heart Disease Data Set [18] donated by David W. Aha. The dataset comprises four databases, namely Cleveland, Hungary, Switzerland, and the VA Long Beach. For this research, two databases namely Hungary and Switzerland are adapted for analyzing disease. The execution of the proposed strategy is performed in JAVA with PC having Windows 10 OS, 2 GB RAM, and Intel i3 core processor.

24.4.1 Comparative Analysis

The methods taken for the assessment include: NN [19], SVNN [20], SVM [13], RideNN [17] and the proposed Rider-DBN.

24.4.1.1 Analysis Considering Hungarian Dataset

Figure 24.2 depicts the comparative analysis of techniques using the Hungarian dataset with accuracy, sensitivity, and specificity. The analysis of techniques with accuracy is deliberated in Fig. 24.2a. When 50% of data is used for training, the accuracy obtained by classical NN, SVNN, SVM, RideNN, and the proposed Rider-DBN are 0.531, 0.605, 0.759, 0.820, and 0.843. The analysis based on sensitivity is deliberated in Fig. 24.2b. The comparative methods, such as NN, SVNN, SVM, RideNN, and the proposed Rider-DBN have the sensitivity of 0.531, 0.605, 0.825,

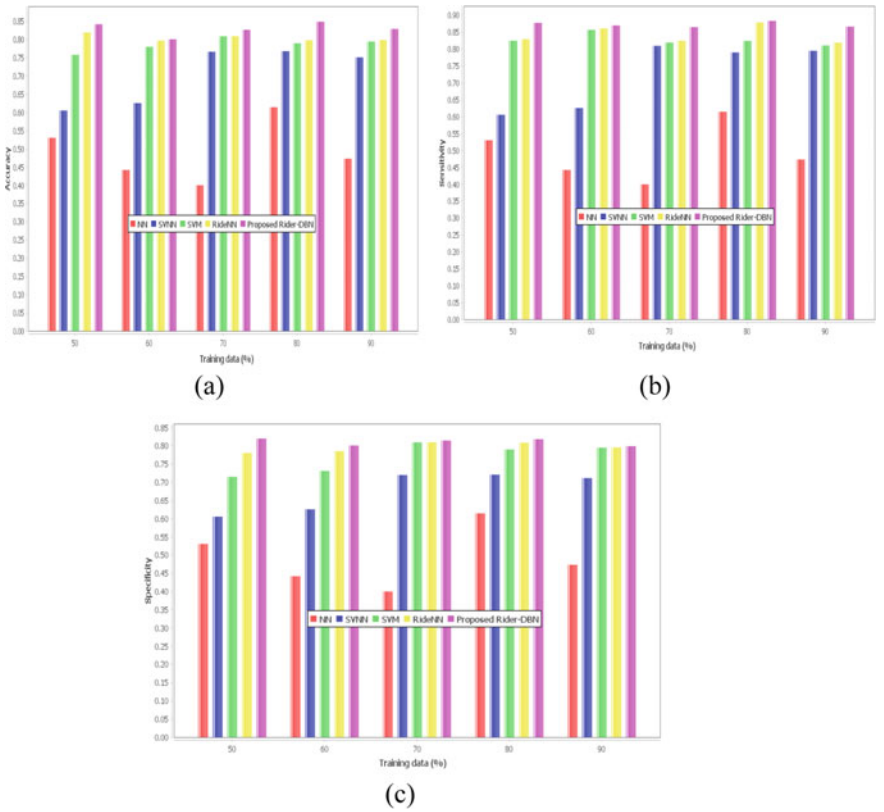


Fig. 24.2 Analysis of methods with Hungarian dataset considering **a** Accuracy **b** Sensitivity **c** Specificity

0.830, and 0.878, when 50% of data is used for training. The analysis using the specificity is deliberated in Fig. 24.2c. When 50% of data is used for training, the specificity of the existing methods, such as NN, SVNN, SVM, RideNN, and the proposed Rider-DBN are 0.531, 0.605, 0.715, 0.780, and 0.820.

24.4.1.2 Analysis Considering Switzerland Dataset

Figure 24.3 displays the comparative analysis using Switzerland dataset. The analysis of techniques using the accuracy is deliberated in Fig. 24.3a. For 50% training data, the accuracy of the existing methods, such as NN, SVNN, SVM, RideNN, and the proposed Rider-DBN are 0.616, 0.830, 0.872, 0.875, and 0.882. The analysis using the sensitivity is deliberated in Fig. 24.3b. For 50% training data, the sensitivity of the comparative methods, such as NN, SVNN, SVM, RideNN, and the proposed Rider-DBN are 0.541, 0.574, 0.683, 0.758, and 0.806. The analysis based on specificity parameter is deliberated in Fig. 24.3c. The methods, such as NN, SVNN, SVM, RideNN, and the proposed Rider-DBN have the specificity of 0.635, 0.893, 0.917, 0.922, and 0.926 when the classifier is trained with 50% of data.

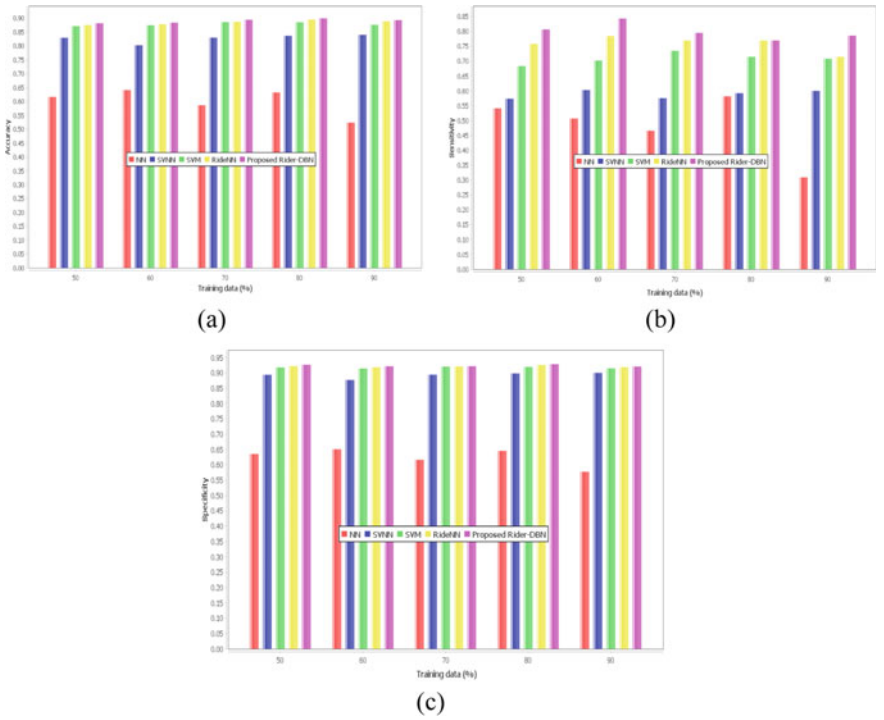


Fig. 24.3 Analysis of methods with Switzerland dataset considering **a** Accuracy **b** Sensitivity **c** Specificity

24.5 Conclusion

This paper presents a Rider-DBN for performing big data classification to predict diseases. The purpose is to provide a MapReduce framework with a deep learning classifier that can classify normal and abnormal classes with better accuracy. The ROA is employed for training the DBN that finds the optimum weights of the DBN for disease prediction. The Rider-DBN provides superior performance with the accuracy of 89.3%, sensitivity of 78.6, and specificity of 92%. In the future, the method will be extended by analyzing the performance of the method with streaming data.

References

1. Tsymbal, A., Bolshakova, N.: Guest editorial introduction to the special section on mining biomedical data. IEEE Trans. Inf. Technol. Biomed. **10**(3), 425–428 (2006)
2. Mane, T.U.: Smart heart disease prediction system using improved K-means and ID3 on big data. In: Proceedings of 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), pp. 239–245. IEEE (2017)
3. Ephzibah, E.P., Sujatha, R.: Big data management with machine learning inscribed by domain knowledge for health care. Int. J. Eng. Technol. **6**(4), 98–102 (2017)
4. Polat, K., Güne, S.: An expert system approach based on principal component analysis and adaptive neuro-fuzzy inference system to diagnosis of diabetes disease. Digit. Signal Process. **17**, 702–710 (2007)
5. Srinivas, K., Rao, G.R., Govardhan, A.: Rough-fuzzy classifer: a system to predict the heart disease by blending two different set theories. Arab J. Sci. Eng. **39**(4), 2857–2868 (2014)
6. Yuvaraj, N., Vivekanandan, P.: An efficient SVM based tumor classification with symmetry non-negative matrix factorization using gene expression data. In: Proceedings of International Conference on Information Communication and Embedded Systems (ICICES), pp. 761–768. IEEE (2013)
7. Vafaie, M., Ataei, M., Koofgar, H.R.: Heart diseases prediction based on ECG signals' classification using a genetic-fuzzy system and dynamical model of ECG signals. Biomed. Signal Process. Control **14**, 291–296 (2014)
8. Long, N.C., Meesad, P., Unger, H.: A highly accurate firefly based algorithm for heart disease prediction. Expert Syst. Appl. **42**(21), 8221–8231 (2015)
9. Kharat, K.D., Kulkarni, P.P., Nagori, M.: Brain tumor classification using neural network based methods. Int. J. Comput. Sci. Inf. **1**(4), 2231–5292 (2012)
10. Han, J., Rodriguez, J.C., Beheshti, M.: Diabetes data analysis and prediction model discovery using rapidminer. In: Proceedings of Second International Conference on Future Generation Communication and Networking, vol. 3. pp. 96–99. IEEE (2008)
11. Lakshmanaprabu, S.K., Shankar, K., Ilayaraja, M., Nasir, A.W., Vijayakumar, V., Chilamkurti, N.: Random forest for big data classification in the internet of things using optimal features. Int. J. Mach. Learn. Cybernet. **10**(10), 2609–2618 (2019)
12 ALzubi, J.A., Bharathikannan, B., Tanwar, S., Manikandan, R., Khanna, A., Thaventhiran, C.: Boosted neural network ensemble classification for lung cancer disease diagnosis. Appl. Soft Comput. **80**, 579–591 (2019)
13. Nalluri, M.R., Kannan, K., Gao, X.Z., Roy, D.S.: Multiobjective hybrid monarch butterfly optimization for imbalanced disease classification problem. Int. J. Mach. Learn. Cybernet. 1–29 (2019)
14. Reddy, G.T., Reddy, M.P.K., Lakshmanna, K., Rajput, D.S., Kaluri, R., Srivastava, G.: Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis. Evol. Intell. 1–12 (2019)

15. Zdravevski, E., Lameski, P., Kulakov, A., Jakimovski, B., Filiposka, S., Trajanov, D.: Feature ranking based on information gain for large classification problems with mapreduce. In: Proceedings of IEEE Trustcom/BigDataSE/ISPA, vol. 2, pp.186-191 (2015).
16. Hinton, G.E., Osindero, S., Teh, Y.: A fast learning algorithm for deep belief nets. Neural Comput. **18**, 1527–1554 (2006)
17. Binu, D., Kariyappa, B.S.: RideNN: a new rider optimization algorithm-based neural network for fault diagnosis in analog circuits. IEEE Trans. Instrum. Meas. 1–25 (2018)
18. Heart Disease Data Set, https://archive.ics.uci.edu/ml/datasets/Heart+Disease. Accessed on April 2020
19. Mathan, K., Kumar, P.M., Panchatcharam, P., Manogaran, G., Varadharajan, R.: A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. Des. Autom. Embed. Syst. **22**(3), 225–242 (2018)
20. Gomathi, N., Karlekar, N.P.: Ontology and hybrid optimization based SVNN for privacy preserved medical data classification in cloud. Int. J. Artif. Intell. Tools **28**(3), 1950009 (2019)

Chapter 25
Modified Compressive Sensing and Differential Operation Based Channel Feedback Scheme for Massive MIMO Systems for 5G Applications



V. Baranidharan, R. Shalini, N. Nithya Shree, S. Sasithra, and R. Jaya Murugan

Abstract In 5G communications, the massive Multi-Input Multi-Output (MIMO) systems are one of the promising key technology for future applications. In the massive MIMO systems, modified compressive sensing and differential operation-based channel feedback scheme is widely used to increase the channel matrix dimension. In this paper, a modified compressive sensing and differential operation-based channel feedback scheme has been proposed to reduce the channel feedback overhead of the massive MIMO systems. In order to achieve stronger sparsity, the temporal correlation-based time-varying channels is widely used and exploited in order to generate the Channel Impulse Response (CIR) between the neighboring time slots. The simulation results show that this modified compressive sensing and differential operation-based channel feedback scheme is reducing the overhead arise due to feedback scheme and achieves the stronger sparsity than the existing systems.

V. Baranidharan · R. Shalini (✉) · N. Nithya Shree · S. Sasithra · R. Jaya Murugan
Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathy, India
e-mail: shalini.ec16@bitsathy.ac.in

N. Nithya Shree
e-mail: nithyashree.ec16@bitsathy.ac.in

S. Sasithra
e-mail: sasithra.ec16@bitsathy.ac.in

R. Jaya Murugan
e-mail: jayamurugan.ec18@bitsathy.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_25

25.1 Introduction

Wireless communication systems will undergo a revolution about once every decade. Now there is a huge demand in the market for a new technology called 5G that should enhance the capacity of the system than 4G technology. The migration towards the higher frequencies leads to increase in the system capacity. For example, millimeter wave (mm wave ranging 30–300 GHz) will release larger bandwidth. The massive Multiple-Input Multiple-Output (MIMO) which contains massive number of antennas (i.e. >= 100 antennas) significantly enhances the spectrum efficiency [1]. The mm-waves are suitable for massive MIMO system because of its higher frequency, very short wavelength, and the physical and enhanced design of the massive antenna arrangement as an array can also be altered.

This method significantly enhances the gain of the MIMO system, and also overcomes the severe path loss of mmWave signals. Therefore, by combining both mm-wave communications and massive MIMO [2], the data rate speed and spectrum efficiency need to be increased. The acquired knowledge about the accurate Channel State Information (CSI) is important in mm wave massive MIMO systems, to improve the link reliability by pre-coding in the uplink and combining in the downlink, because both the Base Station (BS) and Mobile Station (MS) are always equipped with large number of antennas. The CSI channel information acquired is widely used in uplink by leveraging the reciprocity values of the channel. This will be feedback to all the other mobile stations. In comparison with the conventional mmWave MIMO communication systems working at 306 GHz, the channel estimation in the mmWave MIMO systems will be very challenging task due to the large number of antennas, less sparsity, and hardware constraints.

The acquisition of Channel State Information at Transmitter (CSIT) can be done by exploiting the channel reciprocity in Time Division Duplexing (TDD). The problems in TDD like delay in transmitting the information to the receiver due to traffic in the channels are rectified in Frequency Division Duplexing. But still, it fails to hold channel reciprocity which is more important to meet the required system performance of massive MIMO. Conventional codebook based channel feedback method does not suit the massive MIMO systems, because the codebook size needs to be extended in order to acquire a reliable CSIT accuracy [3]. The modified compressive sensing and differential operation-based channel feedback methods based on non-codebook, uses the compressive sensing (CS) to reduce the feedback overhead, since non-codebook possess large channel feedback overhead. CS-based channel feedback methods decrease the overhead by compressing MIMO channel matrix where the spatial correlations are exploited. In practice, when MIMO is distributed, the channels are not spatially correlated. Therefore, the sparsity of time-domain channel is widely exploited to compress the channel impulse response (CIR) directly which leads to the reduction of feedback overhead. But for CIR which is not sufficiently sparse, the overhead is very high.

25.2 Literature Survey

This section explores the many existing techniques which greatly overcome the computational complexity and feedback overhead changes.

Dai et.al. [4] proposed a codebook named 2D space of AoA's and AoD's values, an adaptive subspace to reduce the feedback overhead in FDD multiple access technique-based massive MIMO systems. By using the constant AoD's information, this codebook quantizes the channel vector within the given angle coherence time. But when number of base station antennas is less than 120, the theoretical rate gap is comparatively high.

Sidiropoulos et.al. [5] proposed the Novel Limited Feedback algorithms that lifted the problem of feedback overhead. This algorithm is more suitable for Double Directional Massive MIMO channels where the inherent sparsity is exploited and low computational complexity, particularly on the user side. But, this method is not suitable for conventional uniform dictionaries and the performance gain is poor. Khan et al. [6] proposed a new algorithm called Block Sparsity Adaptive Matching Pursuit (BSAMP). Due to large dimensions of channel vector and unknown value of the channel sparsity, this will lead to computational complexity. Hyukmin and Cho [7] proposed a compressive sensing-based CSI channel feedback method which gains sparsity by exploiting the channels that possess temporal correlation. That is, the sparsity between the channel impulse responses is obtained by temporarily coagulated channels. However, this method results in feedback distortion that further leads to data loss. After analyzing numerically, simulation results show that the sparsity obtained is widely based on the temporally correlated channels values based on the compressed CSI feedbacks.

Kulsoom et al. [8] have proposed a method called Quantized Partially Joint Orthogonal Matching Pursuit (Q-PJOMP). It is also known as Q-PJIHT-based Channel State based algorithm to retrive the various values of CSI from the quantized feedback. In the second stage, Compressed sensing algorithms are used to recover the CSI. The two-step quantization technique is having higher complexity. It also possesses feedback overhead problem.

25.3 System Model

In wireless communication, the information content is sparse due to the random projections. When a massive number of antenna arrays are used, the channel exhibits the sparsity in spatial frequency domain which results in strong spatial correlations and feedback overhead. To overcome this problem, modified compressive sensing (CS) and differential operation based channel feedback scheme is used. Initially, the CS-based method exploits the sparsity by efficiently encoding it without knowing the information about the location. Further, it develops a compression mechanism

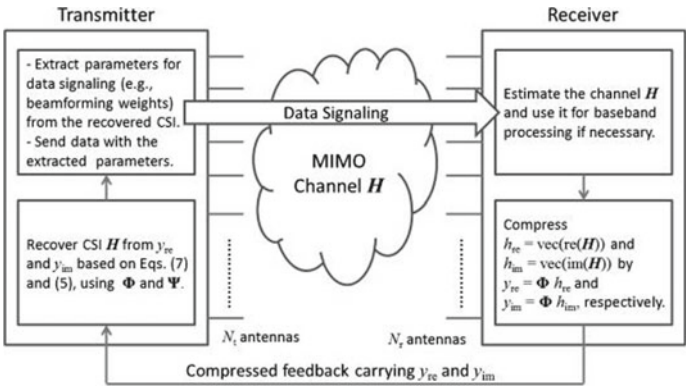


Fig. 25.1 Flow of compressive sensing method

to reduce the Channel State Information by using the small compression ratios (Fig. 25.1).

Consider H as a spatially correlated massive MIMO channel matrix for $N(t) >> 1$ antennas for transmission and $N(r) >= 1$ antennas used for reception where $H = N(r) * N(t)$. In conventional channel feedback methods, the matrix H is directly fed into the channel feedback whereas the CS-based channel feedback method applies a compression mechanism on H before sending it into the channel feedback. It captures the spatially arranged X non zero elements of X sparse with $N*1$ signal vector, where $X \ll N$ by signal will be compressed into $M \ll N$. The compressed feedback content is represented as $Y = \phi * h$.

25.3.1 *Temporal Correlation of Time-Varying MIMO Channels*

The lth element of CIR's support vector $p_n^{(t)}$ is varied over T consecutive time slots and it is given by $\{p_n^{(t)}(l)\}_{t=1}^{T}$. This is known as Markov First Order Process where $p_n^{(t)}(l) \in \{0, 1\}$. The two transition possibilities are represented as:

$$P_{10} = P\{p_n^{(t+1)}(l) = 1 \big| p_n^{(t)}(l) = 0\} \tag{25.1}$$

$$P_{01} = \Pr\{p_n^{(t+1)}(l) = 0 \big| p_n^{(t)}(l) = 1\} \tag{25.2}$$

And the distribution series at time $t = 1$ is denoted as $\mu_n^{(1)} = \Pr\{p_n^{(1)}(l) = 1\}$. If $\Pr\{p_n^{(1)}(l) = 1\} = \mu$ for all time t and n, then $P_{10} = \mu\, P_{01}/(1 - \mu)$. Similarly, the amplitude vector $a_n^{(t)}$ of CIR is varied over time by using the Autoregressive First-Order and it is given as $a_n^{(t)} = \mathcal{P} a_n^{(t-1)} + (1 - \mathcal{P}^2 w^{(t)})^{1/2}$ where $\mathcal{P} = J_0(2\pi f_d \tau)$ and is known as correlation coefficient of Bessel function, f_d represents the maximum

Doppler frequency, τ represents the duration of the time slot and $w^{(t)}$ represents the noise vector which is independent of $a_n^{(t-1)}$. All these parameters are independent and are similarly distributed with zero mean Gaussian distribution $CN\ (0, \sigma_\omega^2)$.

25.3.2 Proposed Modified Differential Channel Feedback Based Compressive Sensing

From the literature survey, many channel feedback reduction methods have been proposed to provide accurate CSI at the transmitter and the receiver side. But the mm-wave massive MIMO antennas are very sensitive to the accuracy of channel state information. A negligible change in the channel state leads to a significant loss in downlink performance, efficiency, and SNR rates. This error level creates a very strong correlation between the successive channel impulse responses (CIR). A good approach to overcome this problem is quantizing the channel impulse response in a differential way because, the time-varying massive MIMO channels have coherence time greater than the time difference between two consecutive channel response.

Differential channel feedback method is a natural method in which the high-level correlation or two-dimensional correlations like temporal or spatial correlation can be exploited to reduce feedback overhead. It uses only the changes in the channel response relative to the previous channel response. That is, it computes the difference between the current and previous channel impulse response by limiting the magnitude of each angle but it maintains the same quantization resolution. Differential quantization is an iterative process. For each iteration, both transmitter and receiver know the accurate information about the current channel response and the last channel estimation. Generally, when conventional feedback methods are used, the receiver feeds back the temporally correlated responses to the transmitter. These responses require greater feedback rate and larger bandwidth. The rate of feedback bit increases when the temporal correlation increases. But the differential channel feedback method reduces the feedback bits to a great extent by exploiting the temporal correlation. Initially, the transmitter predicts the current CIR depending upon the previous CIR. Then, it performs the differential quantization on the channel responses and feeds into the channel feedback. At the receiver side, the quantized responses are captured and recovered. In return, it sends the differential CIR to the transmitter in a compressed way.

The basic idea behind modified compressive sensing and differential operation based channel feedback scheme is that when there is temporal correlation, instead of quantizing the whole channel response which leads to the wastage of feedback resources. The performance of proposed differential channel feedback method is much better than the other conventional methods in low mobility channels. The modified compressive sensing and differential operation based channel feedback scheme differentiates the CIR with stronger sparsity by exploiting the temporal correlation.

The sensing matrix compresses the differential CIR and it is fed into the channel feedback. The received CIR is known to both BS and users in order to recover the original differential CIR by using larger compression ratio. This helps in the avoidance of feedback error propagation because the users execute initialization after certain time slots. The conventional channel feedback method gives feedback about the channel after the estimation in the current time slot estimated values. But the proposed modified compressive sensing and differential operation will be used to generates the sparsity CIR by computing or differentiating two $h(t-1)$ and $h(t)$ estimates where $(t-1)$ denotes the $(n-1)$ which is the previous time slot and t represents the current time slots. Mathematically, the differential CIR $h(t)$ is computed as,

$$\Delta h_n^{(t)} = h_n^{(t)} - h_n^{(t-1)} = p_n^{(t)} \cdot (a_n^{(t)} - a_n^{(t-1)}) + (p_n^{(t)} - p_n^{(t-1)}) \cdot a_n^{(t-1)} \tag{25.3}$$

$$= p_n^{(t)} \cdot [\sqrt{1-\rho^2 w} - (1-\rho) a_n^{(t-1)}] + (p_n^{(t)} - p_n^{(t-1)}) \cdot a_n^{(t-1)} \tag{25.4}$$

Consider ϕ as the Gaussian matrix (sensing matrix) where $\phi\ M * L$ with $M << L$. This sensing matrix ϕ compresses the sparse signal $\triangle h(t)$ into low dimensional vector Y. The low dimensional measurement vector Y is given by $\phi\ \triangle h(t)$ which is fed into the base station. When it is fed into the base station, the low dimensional vector Y is added with the noise in the channel and it is given by the equation $Y = \phi\ \triangle h(t) + n$ where n refers to the noise in the channel feedback. The compressed sparse signal Y is accurately recovered from the noisy base station channel by the most popularly used subspace pursuit (SP) algorithm. This algorithm is chosen for its lesser complexity and noise resistance.

25.4 Simulated Results

The performance metrics are investigated by using the simulation results using MATLAB is explained in this section. The proposed modified CS and Differential operation based channel feedback for massive MIMO systems is simulated and the resulted performance metrics are compared with the existing systems at different channel compression ratio (η). At initially, some parameters are considered for massive MIMO systems for simulations. The simulation parameters used for simulating the system model of massive MIMO systems are tabulated in the Table 25.1.

25.4.1 Channel Impulse Response Comparison

In this method, the channel compression ratio $\eta = M/L$ where, L is the maximum delay spread. Figures 25.2 and 25.3 represents the snapshot of previous Channel

Table 25.1 Simulation parameters

Parameters	Values
Number of the antenna	32
Amplitude vector	$a_n^{(1)}$
Support vector	$p_n^{(1)}$
Channel compression ratio	η

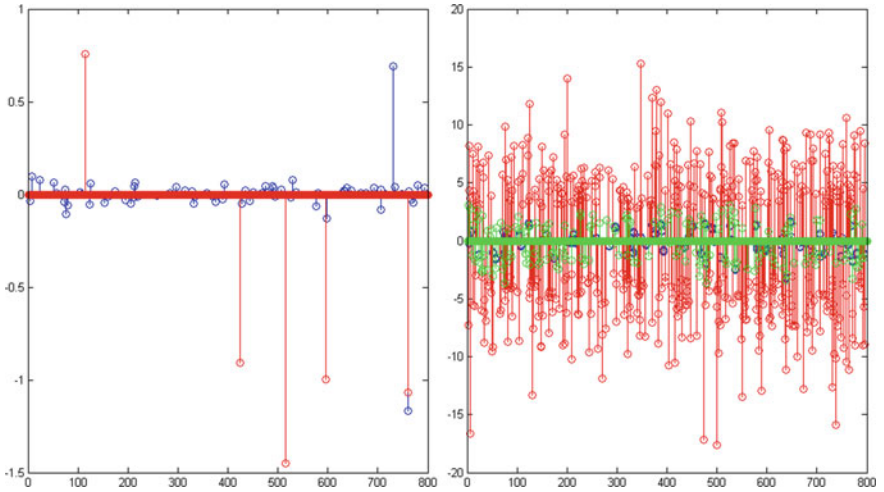


Fig. 25.2 Original CIR

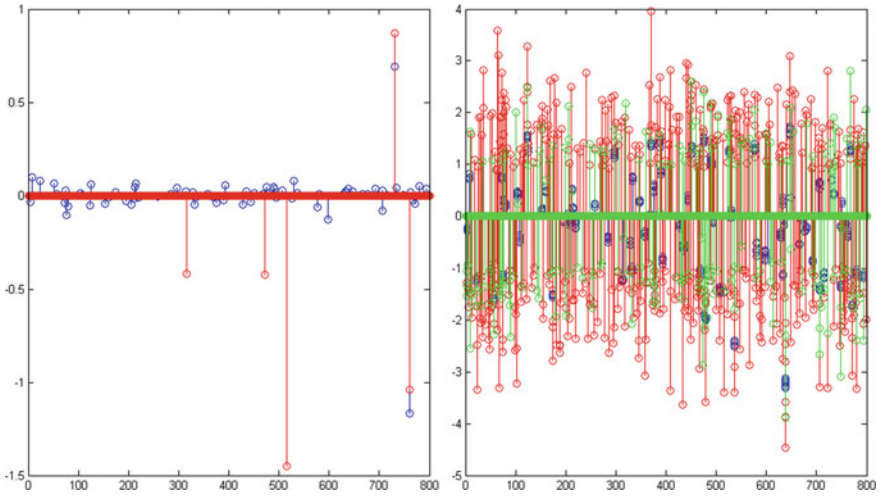


Fig. 25.3 Modified CIR

Table 25.2 CIR simulation parameters

Parameters	Values
L	200
P_{01}	0.05
M	0.1
F_d	10 Hz
T	1 ms
σ_w	1

Impulse Response, current CIR, and differential CIR. The following parameters are considered for simulation and are tabulated in the Table 25.2.

At final, initially the amplitude vector value $a_n^{(1)}$ is considered as random vector. The different entries of independent identically distributed (I.I.D) varied complex Gaussian variables with mean values as zero and unit variance has been considered.

From Fig. 25.3, we observed that the modified differential CIR will give stronger sparsity than the original CIR, which will indicate that it is compressed more but the feedback overhead is reduced comparatively with existing systems.

25.4.2 NMSE Validation

For NMSE validation, initially consider the system with simulation parameters considered for CIR generation. Here, the two different cases are considered $\eta = 25\%$ and the M/L value is changed to $\eta = 45\%$. In the proposed modified differential feedback band channel state information state, η value represents the average range of compression of the CIR over the p feedback slots. The compression ratio of the proposed scheme is 30% for the first case and 40% for the second case. The p-value is considered as 3 for adaptive to the channel condition. Figure 25.4 shows that proposed scheme—based scheme. The direct CS and the proposed differential feedback scheme results statistics are compared and tabulated in the Table 25.3.

In the first case, the existing system work fails to overcome the feedback overhead. But for this case, the proposed system gives good performance. For the second case, the proposed scheme is achieving a better SNR compared with the existing direct CS based scheme. In addition, the proposed scheme for the different channel compression ratio have a very similar NMSE performance, and this scheme can be used to reduce the feedback overhead to ensure the reliable channel feedback.

25.5 Conclusion

In this paper, we analyzed and investigated the issues and challenges in the design of channel feedback with the differential operation in the massive MIMO systems. The

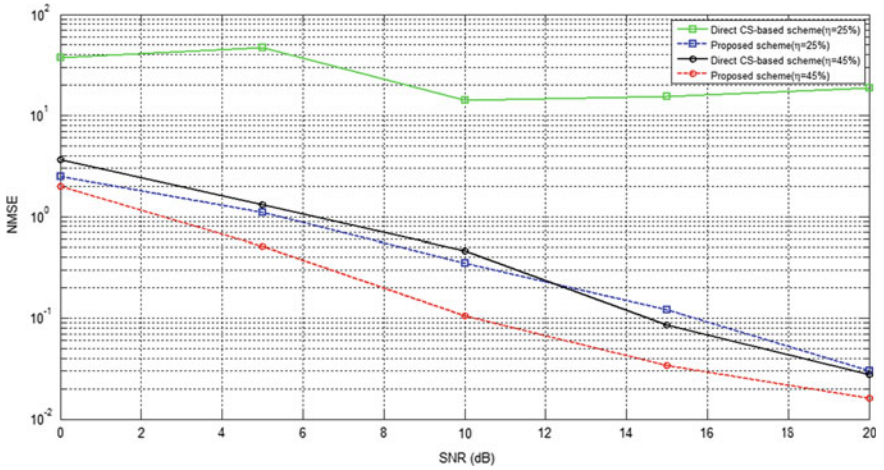


Fig. 25.4 NMSE comparison of the proposed modified differential feedback systems

Table 25.3 Comparison of the proposed differential feedback scheme (CS-DO) with existing systems

Values	$\eta = 25\%$		$\eta = 40\%$	
	Direct CS scheme	CS–DO scheme	Direct CS scheme	CS–DO scheme
Min	14.06	0.03016	0.02743	0.01617
Max	47.31	2.507	3.652	1.999
Mean	26.62	0.8214	1.106	0.5331
Median	18.77	0.3521	0.4642	0.1047
Mode	14.06	0.03016	0.02743	0.01617
Standard deviation	14.93	1.031	1.511	0.844
Angle	33.25	2.477	3.625	1.983

proposed work exploits the modified temporal correlation of the massive MIMO based time-varying channels CIR values and the differential CIR in slots of its neighboring time slots. This modified differential channel feedback scheme gives a stronger sparsity than the existing CIR methods. The modified Compressed sensing and differential operation based channel feedback scheme gives a better Normalized Mean Square Error (NMSE) values than all the other existing methods and it also reduces the feedback overhead for the massive MIMO systems.

References

1. Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive MIMO for next generation wireless systems. IEEE Commun. Mag. **52**(2), (2014)
2. Rappaport, T.S., Xing, Y., MacCartney, G.R., Molisch, A.F., Mellios, E., Zhang, J.: Overview of millimeter wave communications for fifth-generation (5G) wireless networks—with a focus on propagation models. IEEE Trans. Antennas Propag. **65**(12), (2017)
3. Mumtaz, S., Rodriguez, J., Dai, L.: mmWave Massive MIMO, A Paradigm for 5G. Elsevier (2017)
4. Dai, L., Shen, W., Shim, B., Wang, Z., Heath, R.W.: Channel feedback based on AoD-adaptive subspace codebook in FDD massive MIMO systems. IEEE Trans. Commun. **66**(11), (2018)
5. Alevizos, P.N., Fu, X., Sidiropoulos, N.D., Yang, Y., Bletsas, A.: Limited feedback channel estimation in massive MIMO with non-uniform directional dictionaries. IEEE Trans. Signal Process. 66(19), 1 (2018)
6. Khan, I., Singh, M., Singh, D.: Compressive sensing-based sparsity adaptive channel estimation for 5G massive MIMO systems. IEEE Appl. Sci. (2018)
7. Son, H., Cho, Y.-H.: Analysis of compressed CSI feedback in MISO systems. IEEE Wirel. Commun. Lett. 8(6), (2019)
8. Kulsoom, F., Vizziello, A., Chaudhry, H.N., Savazzi, P.: Joint sparse channel recovery with quantized feedback for multi-user massive MIMO system. IEEE Trans. Signal Process. 66(19), 1 (2018)

Chapter 26
Smart Shopping Bag Using IoT



M. Suguna, R. V. Kanimozhi, D. Naveen Kumar, S. Indhushree, and D. Prakash

Abstract Shopping mall is a place where thousands of customers visit every day to purchase many products. Shopping becomes a tedious process. It consumes lots of time as customers have to wait in a long queue. To make shopping easier we have proposed a system titled "SMART SHOPPING BAG USING IOT" where board which will help blind people to verify their purchased product. It will say the details that are displayed in the Liquid Crystal Display (LCD) which will be audible only to the customer carrying the bag. After the final completion, the customer can scan only the bag instead of scanning the product one by one. Once they are done with the scanning, they can pay their bill through mobile phone integrated with Global System for Mobile communication (GSM) module or can pay manually. Products are scanned through Radio Frequency Identification (RFID) tag reader which helps to add up the price of the product purchased. They can switch the mode and remove their product if not needed. This will help to calculate the total amount of their purchase. There is also an encoder-decoder control board which will intimate the customer if they miss their bag or the bags have been replaced. The total value of their purchase is added to their final bill. The bag also features a voice customer are provided with bag which readily displays the product name and details of the item purchased.

M. Suguna (✉) · R. V. Kanimozhi · D. Naveen Kumar · S. Indhushree
Department of Computer Science and Engineering, Kumaraguru College of Technology, Coimbatore, India
e-mail: suguna.m.cse@kct.ac.in

R. V. Kanimozhi
e-mail: kanimozhi.16cs@kct.ac.in

D. Naveen Kumar
e-mail: naveen.16cs@kct.ac.in

S. Indhushree
e-mail: indhushree.16cs@kct.ac.in

D. Prakash
Electrical and Electronics Engineering, VTMT College, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_26

26.1 Introduction

IOT has been emerged as an important technology in our day-to-day life. It is a system of interrelated computing devices, mechanical and digital devices. It has been used to transfer data over network without help of humans. Over the last few years IOT has been used in every field and it make our work easier and faster. Using these new technologies in developing variety of systems led to various edges such as doing away with inconsistency and therefore dependency on human labor, raising accuracy and labor speed. Over the last years, several studies have been carried out based on these premises led to adjustment of both these technologies. IoT has many in-built sensors which has the ability to collect the information and transfer them through the network. This will be helpful in real-life decision-making process. There were a lot of IoT applications which has changed the whole technology revolution. It has its emergence in every field because of its lightweight methodology. One such important field is shopping. Shopping is considered as the main source of economy and the rate at which shopping by people is always increasing at a steady rate. Today shopping malls are considered to be crowded area. During weekend or festival season, the amount of purchasing will be increasing at the peak. Though purchasing things get easier, the main problem arises while billing our purchase. Waiting in a billing section is a time-consuming process. IoT has provided a solution by which the time taken for waiting in a queue has been reduced to 1/3rd of the time spent before. It can make shopping easier by allocating more time in shopping than to billing.

26.2 Literature Review

RFID-based smart shopping kart has an idea of using LIFI technology which is most effective and can minimize the queues. When we enter a shopping mall, its regular to take shopping cart and stop wherever we need the products need to buy and waiting in a long queue for billing the products. Hence it is much time-consuming process. The technology of RFID is used for transmission and photodiode for reception. The product details are collected through RFID. When the products are being collected, we use RFID to transmit the products for billing session. When the shopping cart is taken near the billing part then the products in shopping cart is automatically billed and bill is given to the consumer [1].

Sahare et al. [2] proposed outline of using raspberry PI. It is easier to interface and no need for extra modules to interface with RFID and makes the system size more approachable. The key objective is to deliver low-cost, easily accessible, and a system for supporting shopping. The first step of the customer entering shopping is to take the trolley. Every trolley is attached with a RFID reader. The idea focuses on which the customer inspects the details of the things he wants to buy. While inspecting the RFID tag of the item, a cost of the item is taken and secure in the framework's memory. When the customer scans the RFID tag the information of the

product is passed to the Raspberry pi through serial communication. Raspberry pi will collect the details of ID from the database. LCD will show product details along with the total amount. As soon as purchase gets over key is pressed to transfer the data to the billing counter [2].

Sivagurunathan et al. [3] has explored the idea of using Zigbee technology which directly sends the information to the billing counter. The trolley can automatically guide the customer to the location of the product once they enter the product name. When a customer enters into the mall, they first take the trolley. Every trolley is joined with a scanner tag and a RFID per user. When they purchase an item, the product is scanned through the RFID tag reader and the details of their purchase will be displayed in the LCD placed in the trolley. At the same time, Arduino will pass the information to the PC with the assistance of RS232 extension. Assuming that every single product checked will give a bell sound making the customer ensure that the product is perused. It will be helpful for security reasons. However, tallying is performed as there is no extension of cost particular item in receipt. On extraction of the unwanted product it will automatically get subtracted on their removal. The product information will be shared via Wireless information process and Zigbee transmitter is used for transferring the final billing amount to the billing system. This will improve the security performance and also the speed [3].

IoT applied smart shopping system has visualized the use of RFID-based production data analysis. Thus, the system uses trolley number for the creation of automatic bill. The smart cart attached with RFID reader automatically reads the items. For data processing microcontroller has been installed and for user interface LCD touch screen has been used. Zigbee technology has been used to communicate smart cart with the server because of its low-power consumption and inexpensiveness. Weight scanner is used for weighing items. RFID reader installed before door is used to check that all the items have been paid. Security and privacy-related issues have been considered. Product quality and quantity, customer satisfaction is automatically monitored in this system. Trolley movement has been monitored using Android application which has been implemented in this system. The application is completely based on the Trolley number and total amount of purchased items [4].

Thangavel, Karthick, Karthikeyan, Karunakaran, Prasanth R has proposed as IoT framework where RFID tag is connected to every product, is read by the RFID reader attached to the bag, given per user. Smart racking system can be included, outfitted with RFID per users, which is used to check screen stock, additionally refreshing a focal server. This system uses RFID to check the product and its price. Once the product is detected the system will process and upload the product price and name. If the stock is in demand or the stock is not available, the system will update it in the IOT [5, 6]. This system also detects the expire date of the product and alert the user through buzzer and update in the IOT website. The load sensor will indicate the overload in the trolley. If the system detects overload it will alert the user through buzzer. The LCD display will display the product name, price, and overload [7].

26.3 Purpose and Objective

At present, we are using barcode scanner to scan the products. There is a LCD display which displays the product name and price. The transfer of data from trolley or basket to PC is handled by Zigbee technology. The amount will get added along with the addition of the product. On the completion, the final amount will be displayed in the LCD. It is much a time-consuming process as it more time to get scanned. Customers have to wait in a long queue for paying their bills which is considered to be a tedious process.

The main objective is to build a smart shopping bag which uses RFID tag to scan the product. The details like product name, amount and the expiry date will be displayed. Along with the LCD display speaker is used to say the details displayed. It will be useful for the blind people and senior citizen. Arduino has been used where the codes are embedded.PIC microcontroller has been replaced by Arduino where we can change our codes easily and in an efficient way. Once we add a product the total will get added along with the total. Once the shopping gets over the final total will be displayed and customer can pay the bill either digitally or manually. The digital payment can be made by using customer card using GSM module. Manual payment will be made in the cash counter where customer can show their bag and receive their receipt of shopping.

26.4 Proposed System

We have proposed a system based on recent technology, IOT. Smart shopping bag will be useful for customer in all shopping malls, supermarkets, and dress shop. The bag usually made of RFID tag reader, LCD display unit, speaker and controller board, encoder decoder module. The initial step of the purchase starts with the scanning of RFID tag of the product scanned. The tag will have a unique code which will be displayed in the LCD display along with the details of the product. Ardunio has been used to store the details of the product. The details of the product will be heard through speaker which will be helpful for blind people and senior citizen. The price of the product is added up after every scan of the product. If customer did not need the product anymore, he can usually pull the switch and scan the product to remove that item from the list. The price of that particular product will be reduced automatically. Once the purchase got completed, he/she can press the button which indicates the purchase completion. Customer can pay their bill through the card of their own which is provided by the shop. The total amount and the balance will be pop up to the mobile. During the time of purchase if they miss their bag the encoder and decoder module present in the bag will automatically warn the owner of that bag (Fig. 26.1).

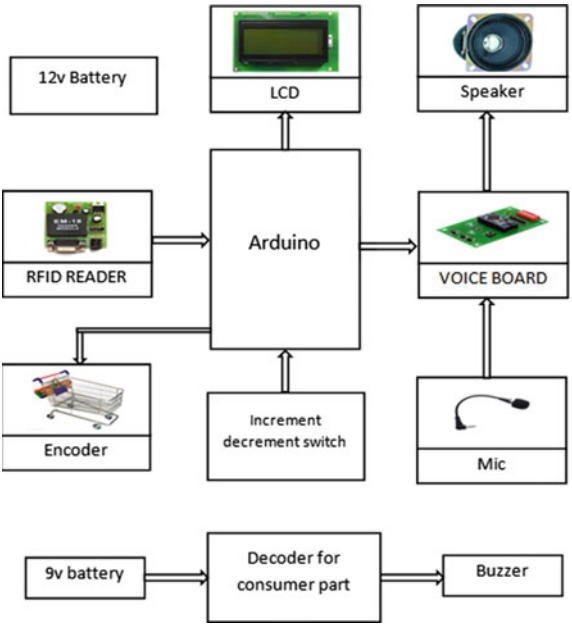


Fig. 26.1 Block diagram of overall structure of smart shopping bag using IoT

26.5 Module Description

26.5.1 RFID Reader Module

This module is usually composed of RFID tag and a reader. The RFID is scanned when the product is placed in the bag using reader which is placed along the bag side. This will be used to calculate the item scanned into the bill. The ID scanned will be unique to every product. This will get matched to the item details stored in the database. Hytertrm software can be used to verify if the scanned RFID is unique or not.

26.5.2 LCD Display Module

The items scanned will be displayed in this module. When sufficient voltage is applied to the electrodes the liquid crystal molecules will be arranged in a way that displays the required information according to the tag they read. The voltage passed should be of +5 V.

26.5.3 Switch Module

This module can be used either for adding or subtracting the cost. The cost of the item that is scanned through bag is automatically get added to their total amount. This will be reflected in the final amount of purchase. If we switch the mode and scan the product the amount of the respective product will get subtracted from the total.

26.5.4 Blind People Module

According to recent survey, 70% of blind people face difficulty while doing shopping. Even though they can identify the product but the price and details to understand are always challenging. The feedback collected from 3–4 departmental stores nearby shows that it will be useful for blind if they had a speaker attached with the basket, so that they can manage on their own without others help. To sort out the problem faced by them, speaker module is developed. WTV-SR comes under the group of recording serial products. WTV-SR module is used for recording and fixing voice playback, recording of the content uploaded and a variety of control modes can be chosen. WTV-SR module is used for various occasions. On the bottom of WTV-SR, I/O can be adjusted by using different control modes. The voice that is recorded already is played while the respective product is scanned. This will be helpful for the blind people who can know the details of the product by voice that are recorded.

26.5.5 Band Module

This module usually composed and encoder and decoder board. This will intimate the customer if they miss their bag or the bag has been replaced. It produces a buzzer like sound when they are moved 5–10 m away from their bag.

26.5.6 Billing Unit

GSM module is used for sending the bill to the respective customer through their registered mobile number. After the purchase has been completed the total will be displayed. If customer pays their bill through card, then the total amount and balance will be displayed in the mobile. The information is transferred through SIM (Subscriber Identity Module) in GSM module to the registered mobile number.

26.6 Results and Discussion

Figure 26.2 displays the Ardunio software where the code is feeded. It is used to check whether the RFID tag is mapped with the correct tag number. It is earlier stage of verification process where the tags are mapped with their respective number.

Figure 26.3 shows the LCD display unit. It displays the name and price of the product which when scanned through RFID reader module attached to the basket. Whenever the item is scanned using reader the details are collected from the database. As soon as the details are retrieved LCD displays the details of the scanned product.

Figure 26.4 displays the integration of various modules. The board usually contains Ardunio, speaker module, voice module, LCD display unit, and a switch. This board is fitted to the basket at the bottom. The main working of smart shopping

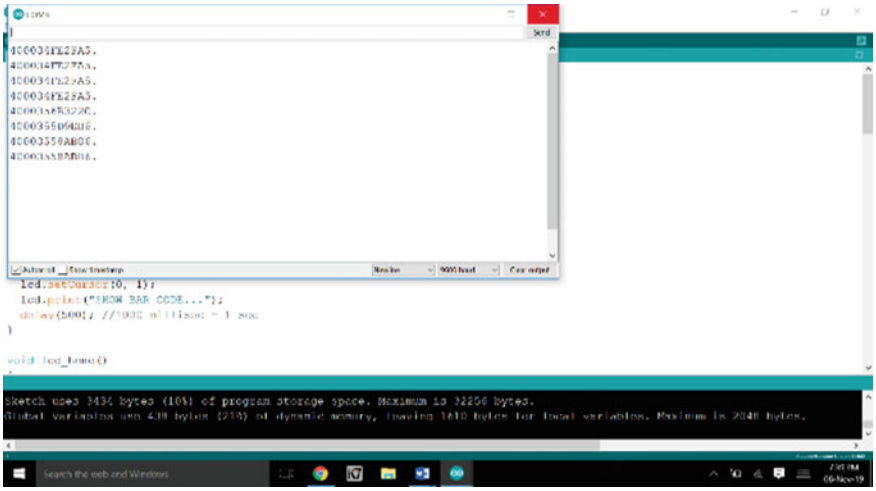


Fig. 26.2 Ardunio verification

Fig. 26.3 LCD display



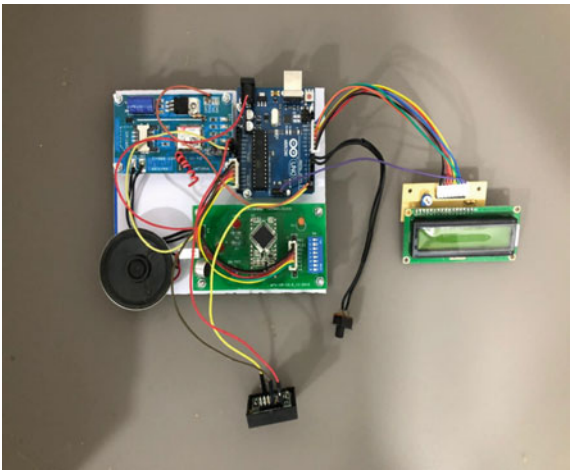


Fig. 26.4 Smart board

takes place in this board which has an interface to various modules. Speaker acts as a mode for blind people to know the details of the product. The speaker in this board is placed to help the blind people for making their shopping better and can be more confident about the items which they buy.

Figure 26.5 displays the page of billing unit. After shopping gets completed customer can use their card to pay their bills. Once the bill is paid then the card balance and the total cost of their purchase will be sent as a SMS to the registered mobile number. The card mentioned is provided by the shop where we can fill the amount in prior and can be used for further purposes.

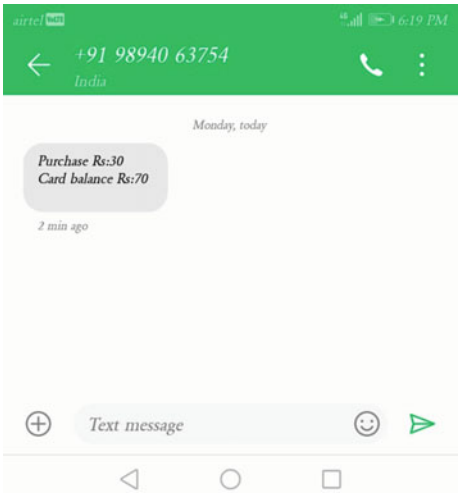


Fig. 26.5 Billing unit

26.7 Conclusion

Thus, the Smart Shopping Bag is made using the recent technologies in IOT. Customer uses this bag for purchase in shopping areas which is kept in the mart so this makes customer easy to purchase products and the total amount will be displayed in the LCD unit. They can scan their bags at the cash counter to pay bills. This will avoid waiting in long queues for billing. Customer can save their time and concentrate more on shopping than waiting in queues for billing. Before billing, they will come to know the total amount of their purchase.

Acknowledgements We affirm that the project titled "SMART SHOPPING USING IOT WITH CUSTOMER ORIENTED SERVICE" is done in focus to help the people who are struggling to do shopping and for waiting in a queue to bill their products. The projects work is carried in our college laboratory and has been developed from the feedbacks received from nearby shopping mart. The information provided is true to the best of our knowledge.

References

1. Sahare, P.S., Gade, A., Rohankar, J.: A review on automated billing for smart shopping system using IOT. Rev. Comput. Eng. Stud. **6**(1), 1–5 (2019). Journal homepage: https://iieta.org/Journals/RCES
2. Sivagurunathan, P.T., Seema, P., Shalini, M., Sindhu, R.: Smart shopping trolley using RFID. Int. J. Pure Appl. Math. **118**(20), 3783–3786 (2018)
3. Vishwas, B., Apoorva, S., Raidurg, S.V., Pawar, A.R., Rananavare, L.B.: IOT application on secure smart shopping system. Int. J. Adv. Res. Comput. Sci. **9**(3), 197 (2018)
4. Devi, R., Pavithira, T., Monisha, E., Likitha, M.: IoT applications on secure smart shopping system. Int. Res. J. Eng. Technol. (IRJET) **05**(03), 1881–1885 (2018)
5. Li, R., Song, T., Capurso, N., Yu, J., Couture, J., Cheng, X.: IoT applications on secure smart shopping system. IEEE Internet Things J. **4**(6), 1945–1954 (2017)
6. Karjol, S., Holla, A.K., Abhilash, C.B.: An IOT based smart shopping cart for smart shopping. In: International Conference on Cognitive Computing and Information Processing, pp. 373–385. Springer, Singapore. Accessed December 2017
7. Shirsath-Nalavade, K.A., Jaiswal, A., Nair, S., Sonawane, G., Suchita, B.A.: IOT based smart shopping cart (SSC) with automated billing and customer relationship management (CRM). Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET) **5**(X), (2017). IC Value: 45.98
8. Shraddha, D.D., Trupti, J.D., Priyanka, S.S.: IOT based intelligent trolley for shopping mall. Int. J. Eng. Dev. Res. **4**(2), 1283–1285 (2016)
9. Li, R., Song, T., Capurso, N., Yu, J., Cheng, X.: Iot applications on secure smart shopping. In: 2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI), pp. 238–243. IEEE. Accessed October 2016
10. Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M.: Internet of things (IoT): a vision, architectural elements, and future directions. Future Gener. Comput. Syst. **29**(7), 1645–1660 (2013)

Chapter 27
An Efficient VM Live Migration Through VM I/O Requests Using Live Scheduling Scheme (LSS)



M. N. Faruk, K. Lakshmi Prasad, and G. Lakshmi Vara Prasad

Abstract Virtual machine (VM) live migration gets essential to execution in the open cloud platform, in order to permit snappy as well as straightforward work deferment scheduling between erratically located hubs, focused on superior and vitality productive use of server framework resources. Regardless, VM Input/Output (I/O) can altogether influence the presentation of VM live storage capacity migration because of conflict aimed at the common disk transfer speed and framework resources. In light of this perception, this paper recommends the Live Scheduling Scheme (LSS) to progress instantaneously the VM execution and migration effectiveness. Live Scheduling Scheme (LSS) lightens the I/O unsettling influence in the middle of VM I/O requests and relocation I/O appeals, by organizing the movement grouping as indicated by the process example of VM I/O demands. Moreover, it advances the store the board productivity by abusing the semantic subtleties from the VMs and relocation threads. The outcomes acquired as of following put together analyses concerning a lightweight model execution of Live Scheduling Scheme (LSS) show that this framework is amazingly robust, enlightening both the VM productivity and migration effectiveness considerably contrasted with prevailing schemes.

M. N. Faruk (✉)
Department of Computer Science and Engineering, Navodaya Institute of Technology, Raichur, India
e-mail: hodcse.nit@navodaya.edu.in

K. L. Prasad
Department of Computer Science and Engineering, Chalapathi Institute of Engineering and Technology, Guntur, India

G. L. V. Prasad
Department of Information Technology, QIS College of Engineering and Technology, Ongole, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_27

27.1 Introduction

In recent years, the expanding selection of distributed computing has changed the IT business. Alongside all-inclusive accessible open cloud arrangements, endeavors likewise exploit comparative arrangements on hybrid and private clouds. Large, virtualized server (Data Centers) is serving the regularly developing interest for storage, networking, and systems administration. The proficient migration of cloud server is progressively significant and complex in the upcoming trends.

Perhaps, web-based administrations, mail management, data source management, and lithography management [1]. The essential hypervisor aimed at the virtualization issue legitimizes the organization of physical devices similarly as the game plan of a collection of cutting edge contraptions to VMs. All the while, hypervisors need to make certain the detachment similarly as reasonableness among diverse VMs on the head of a lone customary physical server, regardless of the way that, improving the basic execution for entire VMs with detectable quality to physical sources [2]. The VM live migration is an action in the part in contemporary hypervisors, that offer assistance to the moving of running VMs starting with one physical agent and continuation within the cloud, either inside the comparable form or all through web server focuses in different regions overall [3].

The key objective of VM real-time migration is to satisfy the boosting essential for lots of coordinating and besides web server commitment mix, system upkeep and update, VM versatility similarly as sensibility in cloud information. With the progression similarly as the colossal use of the VM-based cloud frameworks, in the shared-storage arrangement, VM real-time movement just comprises harmonization of CPU conditions, storage conditions in addition to organize UIs of the objective VM in the middle of resources, and the area of migration [4]. In typical cloud offices, data extra space might be one or the other shared or scattered relying on whether potential data is kept in a focal arrangement, whichever web servers share the physical extra space or scattered (No-Sharing) arrangement, just as each web server consumes its private special submitted extra space [5]. The VM live extra space migration has come to be a basic segment of VM real-time migration in contemporary hypervisors [6]. By way of the executing VM is moving from one hub to multiple hubs, together with the VM's in memory conditions just as computerized disk images, move from the resources web server toward the remote web server. VM online migration, as a crucial down to basic component of the hypervisors, as, QEMU-KVM [7], HyperV [8], and ESX [9], is a much more noticeable advancement. To be point by point, any sort of commotion just as ideal VM live extra space migration framework need to have the holding fast to private or business properties.

(A) *High Performance of VM I/O*: The cloud organization should give precisely the same effectiveness confirmation for each one of those co-found and synchronous executing VMs in the physical web server. All through the VM live extra space migration, solicitations inside the moving VM are as yet executing just as they should be uninformed of the migration methodology. Offered extra space source is radically reached out due to included resources

hungry migration threads. The I/O proficiency inside running VMs needs to keep the Service-Level Agreement (SLA) in any capacity times.

(B) *Proficiency to Migrate Several VMs Concurrently*: It is significantly harder to achieve the suitable VM I/O effectiveness of all running VMs as opposed to relocating various VMs to their areas at a reasonable migration rate. Given the huge arrival of VMs in data offices, it wins to move a few VMs further than a multiple web server or move a few VMs directly obsessed by one singular web server.

(C) *Less time utilization on migration*: The component of a VM's extra space picture is much of the time various to 10 s of GBs just as it would take mins or additionally hrs to comprehensive a VM live extra space migration that is well on the way to diminish the limits of framework checking in the cloud data. In contemporary data offices, VMs are running every minute of an everyday premise to offer purchasers around the world.

An assortment of strategies has been prescribed to support the real-time extra space migration proficiency [10], comprising of enhancement on the data impede transmission arrangement and the migration procedure, limiting tedious data transmission, and utilizing heterogeneous capacity gadgets. Because of these checking, this paper suggests the Live Scheduling Scheme (LSS), that timetables the block migration arrangement taking into consideration of VM I/O requirements, to guarantee the burdening disk head exercises could be brought down considerably [11]. On the above of that, by interestingly storing data chunks after movement threads, inbound VM I/O requirements would be offered in the memory, along these lines reducing the assortment of memory gets to all the while. By doing this, both VM I/O productivity and migration proficiency can be supported extensively.

27.1.1 Definitions

Sequential I/O attribute: The productivity of Sequential I/Os is far superior to that of Random I/Os with standard disk searches for and revolutions [26]. The continuous I/O private property is recognized by loads of measurements from both spatial just as temporal dimensions [27] Consecutive I/O building has been only one of the best central standards in framework exploration study area [28], by and large, because of the proficiency variety in the middle of Sequential I/Os just as Random I/Os away storage space frameworks. *Temporal Dimension*: This is the middle of progressive I/O requests. When there is an extended holding up period over the middle of two I/O requests, a couple of different accounts I/O requests could be given in the middle of, which will unquestionably impact the successive I/O private or business property of the underlying I/O stream. Interleaved Streams. This will resemble a blend of I/O requests commencing a few threads, solicitations, or virtual environments. *Spatial Dimension*: The component of statistics to look at or form in the I/O request commonly. Progressive Addresses. Nevertheless, the qualification among LBA of

progressive I/O requests remains inside a previously classified limit. Perhaps more classified as simply progressive just as stride access.

A few streamlining techniques have been made to improve the back to back I/O private or business stuff by utilizing the relative tips from the applications. The proficiency of the Random Write I/O stream processed within and it will positively besides be weakened impressively. Different examples of productivity decay for various interleaving I/O streams are situated in [29]. In the log-based records framework [30], enough assortment of modernizes is cushioned in memory previously they remain conveyed to disks as an immense back to back part request to ensure that the disk throughput is enhanced drastically contrasted with private minimal subjective make requests. Look at requests are offered in a practically identical manner, which will positively audit an all-out area from disks simultaneously. Offered the idea of disk openness characteristics and back to back I/O origin of interleaving I/O streams, current information frameworks, both local reports frameworks [31] just as scattered information frameworks [32], support the I/O effectiveness by emphatically conveying sequential I/O requests to the hidden hard disk drive.

27.2 Related Works

The virtual machine (VM) live migration is extremely famous through mutual storage condition, where the two sources just as goal servers protect access to the basic extra space framework. Other than the exchange of a determination of VM, for example, CPU, system, memory and physical devices, moving VM's memory conditions regularly proceeds the greatest bit of relocation time just as an assortment of procedures are anticipated to build the memory migration management represented by [12].

A remarkable procedure CR/TR-Motion [13] is recommended to move checkpoints just as execution follow records instead of memory pages at the pre-copy stage. Furthermore, the author [14] utilizes remote direct memory access (RDMA) to move VM relocation traffic. The post-duplicate strategy, some time ago investigated with regard to process relocation abstract works, has been acquainted with move VMs [15]. The memory "copy" period of the live development is conceded until the VM's CPU conditions have been moved to the objective hub. Pre-duplicate is the commanding ongoing relocation procedure to do the live movement of VMs. These comprise of hypervisor-based methodologies, for example, VMware [16], Xen [17], and KVM [18], working framework degree procedures, perhaps OpenVZ framework [19], just as development over the wide-zone organize [20] Every one of the above frameworks utilizes the pre-duplicate calculation for live VM relocation in a memory-to-memory strategy.

Author [21] identified the downtime of pre-copy migration by the writable operational collection (WWS) of the solicitation. The WWS is a collection of web pages being regularly upgraded to a scope that it is risky to transmit them earlier the last iteration. Subsequently, it will certainly be upgraded once again in every single short duration, as well as any previous transmissions would be lost. Nonetheless, the author

did not examine the association between WWS as well as the network condition. A well ahead piece of work by [22] offers simulation designs which can forecast the complete movement time and also interruption (downtime) depend on the pre-copy arrangement. There are two or three examinations considers that have tended to the capacity development of VMs. Authors [23] and [24] applied for Xen VMs through backend disk driver, which can move disk images of a running VM to an area have, in corresponding with memory replicating. Until the third migration of movement depicted in Area II, a moving VM does all I/Os through a source has, making fresh out of the box new squares to the disk images. All through the live development, the backend driver copies all disk squares and more than once moves refreshed squares during disk relocation. An advantage of this component is that there is no exhibition pulverization after development is done. It is once in a while hard to change having hubs quickly; enormous disk development needs long exchange time, in spite of the fact that that must be finished before VM reactivate on an area have. Author [25] uses a duplicate on-compose disk gadget made out of topic images and overhauled squares. A layout disk image, that incorporates standard framework programs that are basic to all VMs, is duplicated to all host hubs ahead of time.

27.3 Proposed Method

This segment focuses on Live Scheduling Scheme (LSS) and its work principle. The second segment explains the block migration process and its scheduling using LSS cache manager. The design objectives as follows. With the intention of progress the enactment of VM live migration, the time constraints play an important role in disk-level operations. The VM I/O performance can be improved via block cache hit ratio and significantly decreasing the disk seek setups. Offering sufficient flexibility with our proposed model, the proposed model flexible enough with respect to diverse applications can be achieved through cache replacement procedures and flexible movement of chunk sizes.

Document-based computerized disk images have been altogether taken on in the virtualized setting [5]. Commencing the VM's perspective, the web-based virtual images concur as to the logical disk that continues all courses of the block layer's application program interface, for example, iSCSI procedural instructions. Likewise, it works with the mainstream remote OS for instance Windows and Linux OS's. Commencing the extra storage framework's perspective, web-based virtual disk images remain considered as tremendous just as expected reports that can be kept in a great deal of records frameworks. The whole thing of I/O requests since client uses of in succession VM will change the I/O requests aimed at hidden enormous reports that embrace advanced disk-based ISO images. The development stream remains positively audit these colossal records to move the VM's extra storage conditions subtleties. In the direction of support the productivity of online disk images, various submitted document/storage frameworks have been made for advanced ISO images only, for instance, VMWare's VMFS and VM Store. LSS is a fundamental

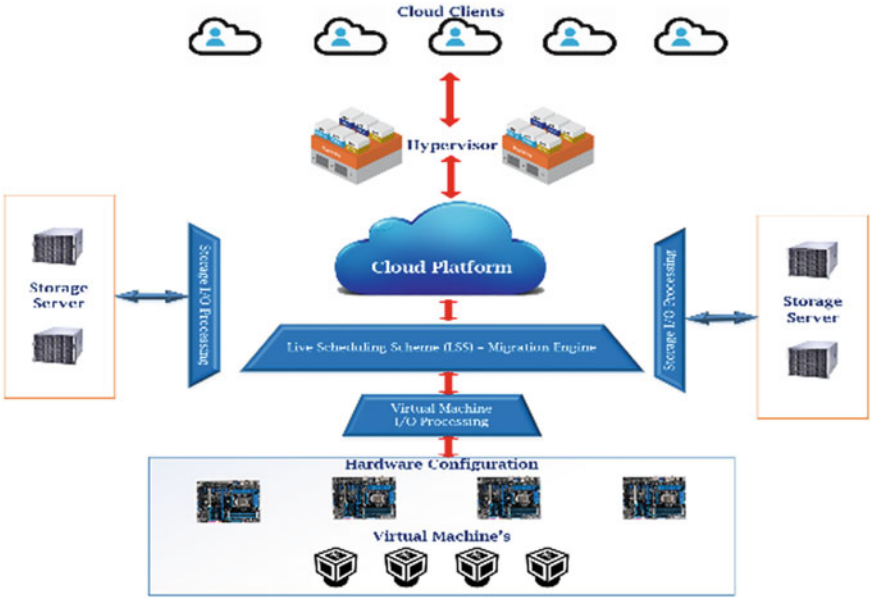


Fig. 27.1 Live scheduling scheme (LSS) system model

yet manageable part that can be integrated directly into present-day hypervisors, for example, XEN [8] hypervisors, and ESX [9]. The aforementioned determinations as development piece measurement and chunks reserve substitute could be tweaked to diverse application work. Aimed at the VM live extra storage development work, merely the web server in the resources side requires to combine the LSS segment, although the webserver on the area side remains by no alterations. LSS is an effectiveness increment layer that could be joined with a standard online development technique, comprising of Dirty Block Tracking also furthermore IO Mirroring, toward extra upgrade both VM IO proficiency just as development productivity. LSS can Fig. 27.1. The structure of the LSS System is put on live to move VMs in the middle of web servers inside exactly the same assortment or all through data offices in diverse territories.

27.3.1 I/O Threads Level Interference

At the point when it concerns the VM live extra space development, one passing thread will surely be allocated for every single moving VM, and also furthermore, it would unquestionably migrate overall conditions evidence of the migrating VMs since the resources web server toward the area web server. Subsequent off, for every executing VM, or else single VM-level I/O stream is made through the interleaving

of various application-level I/O streams by means of similarity VM Operational framework I/O stream. From the perspective of the Sequential I/O private attribute, the passing thread is an astounding successive work, as it will absolutely look at data inside the space of the computerized disk photographs consecutively and ought to have incredible proficiency. That isn't the reality. All through the VM live extra space development, hypervisors are as of now overpowered as a consequence of the surplus data transfer capacity hungry passing threads. By all the more destroying the Sequential I/O private or business resources of passing threads, the further weight will unquestionably be situated on the development motor, to guarantee that is far harder to move VMs rapidly just as flexibly SLA for the I/O proficiency of all getting included VMs. Contemplating all measurements associated with the Sequential I/O building, VM-level I/O Streams (called as VM I/Os) are fundamentally settled by client applications just as remote operating systems. Consider the following Fig. 27.1 uncovers the aggravation between I/O threads in a virtualized framework all through diverse stages of VM live extra space development. The interleaving of entire VM-level I/O streams in addition to the hypervisor I/O stream would positively come to be the keep going I/O stream for the extra space framework.

27.3.2 System Execution and Workflow

Investigating carefully the cutting edge of related examination study employments just as development tasks inside frameworks, we have twin monitoring's: (A) Specific I/O requests commence the roaming VM is created by solicitations, to ensure that we consume restricted abilities to change the VM I/O stream to support the VM I/O effectiveness; (B) Only the general development time just as the precision of the VM conditions broadcast issues for passing threads, while the particular interest measurement, the preliminary location of every single appeal or the arrangement of development I/O requirements remains not comprised of as main concern ones. Consequently, the proposed framework will unquestionably help both the VM I/O proficiency just as development effectiveness by creating and setting up the development arrangement as indicated by the I/O stream of VM requirements. Fundamentally, this resolves positively pick the accompanying data block development prospect dependent on dual prerequisites: (A) This data block has not to remain moved to the area; (B) the concerned location of this data-center is near the contemporary openness region or the situation of the disk head. The anticipated work can accurately reserve in memory data deters looked at bypassing threads from the extra space framework, just as utilize the above data to offer the determined inbound VM I/O requirements, with the goal that the above VM I/O appeals resolve unquestionably not require to get to the extra space storage framework by any stretch of the imagination. Both the VM I/O productivity just as development I/O effectiveness can be improved drastically. In existing virtualized frameworks, there are dual-level cache-store intended for every single running VM: remote ISO makes cache reserve (inside the VM) in addition to host site page cache reserve (inside the hypervisor). LSS has two key segments:

Scheduler for Block Migration as well as Cache Management Layer. Scheduler for Block Migration will surely assess the VM I/O requirements web traffic, perceive the existing VM access the zone, envision the far ahead I/O availability segment, and after that pick the best information part to move. When the information partition is moved to the area web server, it will unquestionably be gone over to the Cache Management Layer. The choice of development information pieces depends on dual divisions: A. a lot shorter search for the season of the extra space framework for the development I/O request; B. the current information block brought by the passing thread might offer well ahead on VM I/O requests with a more prominent chance. Cache Management Layer is to deal with the memory source and carefully cache information deters for advanced VM I/O requests. For instance, the passing thread just checks the computerized disk ISO images when it determines unquestionably not get to exactly the equivalent information deters extra noteworthy than when, other than the redesigned information chunks. These information chunks can be gotten to by VM I/O requirements, just as thusly, as reserving warm relocation information hinders in memory, a few inbound VM I/O requests could be offered by Cache Management Layer in memory straight. When the block cache is finished, a cache substitute formula resolves positively go about just as cool information chunks will surely be expelled.

This is an old as well as typical issue in several facets of system layout, which call for various methods, such as on the Internet profiling as well as trace evaluation, to tackle this trouble. Vibrant piece dimension has been used to lower the number of duplicated information portion movement. In the suggested LSS system, we begin with taken care of piece dimension as well as review its efficiency enhancement. There is no question that LSS can be incorporated with various other piece dimension resolution devices for various work. The complete formula of Migration Chunks Scheduling exists in Algorithm 1.

The very first choice in which system managers require to make is just how much storage space sources need to be designated for the movement string. Provided the complete storage space source continues to be the same, the extra source movement string obtains, and also the much less storage space data transfer VM threads have. Basically, it been compromise in among VM I/O efficiency and also movement efficiency, as well as it can be readjusted for various usage situations or choices. As quickly the storage space source allotment is established, LSS scheduler for Block Migration will certainly pick the following portion to check out from the storage space system; however, this portion can certainly be moved to the location web server. Preferably, we like to information portions pleasing the list below problems: (A) The portion has not remained moved to the location web server yet, (B) The beginning address of this information piece is adjacent to existing VM I/O gain access to area, (C) High chances that the inbound VMI/O demands will certainly review partial or loaded with this information portion.

Algorithm :: 1
Procedure :: LSS scheduler for Block Migration

Step –1: Process initiation
Step –2: DefVMMigration – Describe the group of presently migrating VMs
Step –3: CurrLocation – The storage counter location
Step –4: Func(I/OSchedule (DefVMMigration)) {
Step –5: for (Present VM(1) in the DefVMMigration) {
Step –6: Attain first I/O request for the present VM(1);
Step –7: Update:
Step –8: Current spot with respect to I/O request and its present CurAddress and chunk sizes
Step –9: AvailableVMs – The overall chunks available for VM to migrate.
Step –10: PickSelectedChunk(AvailableVMs, CurrLocation)
Step –11: }
Step –12: }
Step –13: Func(PickSelectedChunk(AvailableVMs, CurrLocation)) {
Step –14: if (AvailableVMs != empty) {
Step –15: pick data block from AvailableVMs based on CurAddress and ignore CurrLocation
Step –16: Deselect data block from AvailableVMs
Step –17: Dissipate VM I/O request for the present data block
Step –18: Update:
Step –19: Current spot with respect to I/O request and its present CurAddress and chunk sizes
Step –20: else
Step –21: Terminate the loop }}

Cache Management Layer all through the runtime, respectively, executing VM is designated an assortment of memory information through the hypervisor, to guarantee that the VM could cache subtleties information in memory, moderately openness of extra space device. Typically, this memory site information is distinguished as compose make disk cache and has website page cache, as clarified in the past segment. As it concerns the VM live extra space development, the entire online disk ISO images will unquestionably peruse from the extra space framework to memory and after that moved to the storage web server. A request follows: in circumstances, we don't cache altogether data impedes in memory in any capacity, we transpire discarding an extraordinary segment of extra space information migration. Thinking about the value it requires to look at in data hinders from the extra space framework to memory, just as included read requests, can be surely delivered to the extra space framework by VM thread, likewise, stipulate the objective data chunks consume very perused in when by the development thread. Reserving an assortment of data parcels with a more noteworthy chance to be gotten to by VM thread in memory would essentially support the VMIO effectiveness by the diminishing of extra space availabilities. The hindrance is actually how to perceive these data chunks just as exactly how to inspect just as change data partitions as soon as the cache is finished. For instance evaluated at past region, the Migration Chunks Scheduling part composes the absolute first snag, while the Cache Management Layer settle in the subsequent one.

Algorithm :: 2
Procedure :: *Cache Management Layer*

Step –1: Process initiation
Step –2: Connect to hypervisor and map memory information
Step –3: Process initiation for metadata of the Cache management layer
Step –4: Collect the request from the respective threads or currently migrated threads
Step –5: Func SetupBlock(AddressBlock){
Step –6: if (AddressBlock >= Cache Management Layer)
Step –7: { Update the current block and the counter value becomes zero
Step –8: else {
Step –9: Setup the block from storage space and dissipate block to end-user
Step –10: Manually migrate the current block to the respective allocation
Step –11: }}}
Step –12: Func UpdateBlock(AddressBlock, Buffer) {
Step –13: if (AddressBlock >= Cache Management Layer)
Step –14: { Update the current block and the counter value becomes zero
Step –15: else {
Step –16: Update the block from storage space and dissipate block to end-user
Step –17: Manually migrate the current block to the respective allocation
Step –18: }}}
Step –19: Func DisableBlock(AddressBlock, Buffer){
Step –20: if (AddressBlock < Cache Management Layer)
Step –21: { Disable the current block and the counter value becomes zero
Step –22: else {
Step –23: Disable the block from storage space and dissipate block to end-user
Step –24: Manually remove the current block to the respective allocation
Step –25: Update the status of Cache management layer
Step –26: Terminate the loop }}}

In the Cache Management Layer, each entrance is a file segment of an online disk ISO image in a precise location, and it includes an assortment of respective access in the Cache Management Layer. For every single passageway, we can survey its liveness worth dependent on the answer for agreeing to concerns: (1) If the data block has very be situated or moved to the area web server, (2) Overall period of this data block stayed inside the cache, (3) Finally, the open door that this statistics block can be unquestionably gotten through VM IO thread henceforward. Based on the respective data, Cache Management Layer can orchestrate these passageways just as change the entrance with the least present worth as the cache is finished. Given that such material evidence is intently relating to the spatial just as the fleeting region of work, such a cache checking equation should be adjusted to give diverse applications. The significantly additional exact district we gain from work, additionally improved cache struck extent and IO productivity that would attain. Differentiated to the standard method, in which the normal 2-degree cache organization equation is utilized, the proposed migration-mindful cache checking framework can impressively help the VM live extra space development proficiency. The skeletal framework equation of the Cache Management Layer is outlined in Algorithm 2.

27.4 Experimental Analysis and Discussions

In this segment, we give the proficiency assessment of the LSS conspire to employ broad follow driven trials. As talked about already, both VM effectiveness and VM constant movement proficiency are essentially imperative for generally system productivity. In the theoretical system, our study concentrates on the assessments of capacity productivity for mutual threads of VM as well as migration. Especially, our version establishes a virtual hard drive image in the hard disk, equally replay the supplied storing chunks level observes on the head of the virtual disk image. Aimed at together VM IO efficiency equally as relocation implementation, our scheme makes use of standard IO critical remarks for efficiency measurements. Shorter feedback time for VM I/O asks for infers additional IO implementation for the modifications inside the running VM.

The designed LSS program is going to be compared to the high-quality transfer scheme which will definitely overlook the features of the VM IO flow in addition to shift the digital hard drive images as of the start to the termination, like the standard VM transfer in Linux Kernel Virtual Machine, the traditional conditions suggest that the structure is actually executing deprived of VM movement. All through the course of investigating, numerous specifications are thought about. Initially, the present migration portion measurements are taken into account within the practices, and also for that reason, the LSS unit performance is usually more improved focusing on innovative vibrant chunk dimension movement strategies. Second, our system allocates the storage space resource in between VM threads as well as transfer threads with a pre-determined proportion. This technique is straightforward but reliable to discover a compromise in between the VM functionality and also migration efficiency. The recommended LSS unit likewise can easily strengthen the VM stay storing migration functionality under various other storing information allocation plans. The LSS unit is analyzed under the various amounts of simultaneous VM reside live storage space migrations.

The investigational system contains a server set up with one AMD Ryzen 3000 series processor, 16 GB DDR4 RAM and double 2 TB disk drives, 12.10 Ubuntu OS. With the intention of evaluating the performance remodeling of the LSS scheme, our system evaluates chunk-level traces as well as likewise pick up IO performance appropriate details throughout the course of the migration plan. To explore exactly just how the movement chunk size, the storing distribution strategy in addition to the selection of simultaneous VM movement impact the enactment of the LSS model, trace-driven trials are executed. The standardized action opportunity of the LSS plan depend on the basic system under the same trial arrangement is studied. Prior to the beginning of the VM remains storage space movement, the policy of keeping source appropriation in between VM IO thread along with transfer thread is developed either through system admins or instantly. As a variety of consideration of VMs shares a singular web server, as well as the organizing server, can easily not please the criteria of storage space IO data transfers for each specific VMs, one or more VMs will most

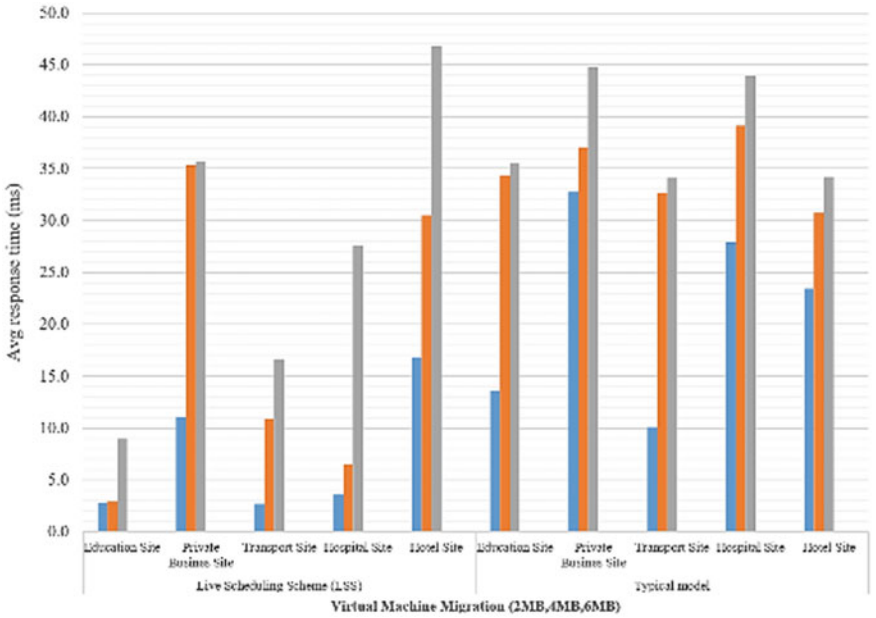


Fig. 27.2 Performance comparison of both typical and LSS model

definitely require to come to be online moved to different other organizing web servers.

The migration effectiveness and VM execution in the middle of LSS and typical movement strategy remain inspected and differentiated for follows beneath diverse migration piece measurements (2, 4, 6, 8,10, and 12 MB) is referenced in the Fig. 27.2. The IO migration time in LSS is balanced out dependent on that in the fundamental migration. In this way, LSS can support the VM live storage space relocation proficiency below various movement divide measurements. As investigated in past areas, a few simultaneous VM live storage space movement wins in the current central CPU structure. So as to explore the amount LSS plan can help the VM live storage space migration execution contrasted with regular development framework in these circumstances, we perform tests to look at productivity enhancement under diverse follows as also diverse simultaneous VM live storage movement (Fig. 27.3).

On this specific occasion, there is a solitary VM moving originating from the source to the area. The migration chuck estimations are 2 MB, and the extra space source allocation proportion is partial between VM and migration threads. The proposed adaptation worked out with diverse traces, contrasted with the ordinary migration framework. To guarantee that excessive find just as pivot capacities are decreased unmistakably in the move I/O requests. This observation makes less confounded for the hard drive administrator to interest inward advancements, for example, the requests converge, with the intention of extra improvise the IO proficiency. Thusly, the general development of IO proficiency is fortified impressively.

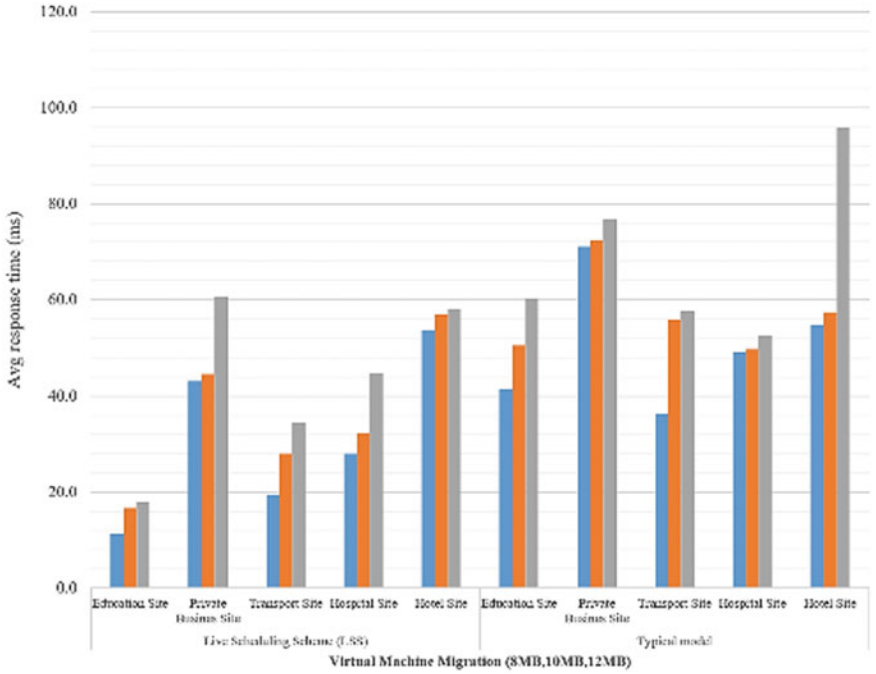


Fig. 27.3 Performance comparison of both typical and LSS model

The explanation liable for such upgrade is really that the incorporated IO execution torrent at the door of extra space is coming to be considerably more back to back in the LSS framework than that in the typical framework. The access to the zone of the applications can without much of a stretch be reserved.

27.5 Conclusion

In view of these above interpretations, this paper provides LSS, a unique VM live storage space movement plan that enhances equally the VM I/O efficiency as well as migration efficiency by producing and also arranging migration structure conferring to the I/O stream of VM demands. By doing this, it is predictable to lessen as well as do away with excessive disk head motions; consequently, the common VM live storage space movement efficiency could considerably be enhanced. Examinations of trace-dependent investigate showed that LSS perhaps minimize the ordinary I/O reaction period via 67% for migration thread as well as 49% for VM I/O string, contrasted to the conventional movement. Subsequently, numerous unneeded taxing disk head looks for as well as rotation procedures that degrade both the VM I/O efficiency and migration enactment were familiarized. Present VM live storage space movement

methodologies neglect the sequential I/O residential or commercial resources of the interleaving I/O streams at eviction of the storage web server, by just transferring the VM's online disk images successively, no matter the simultaneous VM I/O streams too various other migration streaming methods.

References

1. Faruk, M.N., Sivakumar, D.: Multi-layer QOS based task scheduling algorithm for cloud environments. Int. J. Adv. Comput. Technol. (2014)
2. DeHaan, M.P.: Methods and systems for cloud management using multiple cloud management schemes to allow communication between independently controlled clouds. US Patent 8,271,653, 18 Sept 2012
3. Srinivasan, D., Wang, Z., Jiang, X., Xu, D.: Process out-grafting: an efficient "out-of-VM" approach for fine-grained process execution monitoring. In: Proceedings of the 18th ACM Conference on Computer and Communications Security, pp. 363–374 (October 2011)
4. Checconi, F., Cucinotta, T., Stein, M.: Real-time issues in live migration of virtual machines. In: European Conference on Parallel Processing, pp. 454–466. Springer, Berlin, Heidelberg (August 2009).
5. Mao, B., et al.: IOFollow: Improving the performance of VM live storage migration with IO following in the cloud. Future Gener. Comput. Syst. **91**, 167–176 (2019)
6. Beveridge, D.J.: Coordinated hypervisor staging of I/O data for storage devices on external cache devices. US Patent 9,081,686, 14 July 2015
7. Chandak, A., et al.: Dynamic load balancing of virtual machines using QEMU-KVM. Int. J. Comput. Appl. **46**(6), 10–14 (2012)
8. Pate, S.D., et al.: System and method for secure storage of virtual machines. US Patent 9,053,339, 9 June 2015
9. Waldspurger, C.A.: Memory resource management in VMware ESX server. ACM SIGOPS Oper. Syst. Rev. **36**(S1), 181–194 (2002)
10. Van Tran, TK, et al.: Size estimation of cloud migration projects with cloud migration point (CMP). In: 2011 International Symposium on Empirical Software Engineering and Measurement. IEEE (2011)
11. Duggan, M., et al.: A multitime-steps-ahead prediction approach for scheduling live migration in cloud data centers. Softw. Pract. Exp. **49**(4), 617–639 (2019)
12. Hines, M.R., Deshpande, U., Gopalan, K.: Post-copy live migration of virtual machines. ACM SIGOPS Oper. Syst. Rev. **43**(3), 14–26 (2009). Nelson, M., Gailly, J.-L.: The Data Compression Book, 2nd edn. M & T Books, (1995)
13. Hines, M.R., Gopalan, K.: Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning. In: Proceedings of the ACM/Usenix International Conference on Virtual Execution Environments (VEE'09), pp. 51–60 (2009)
14. Huang, W., Gao, Q., Liu, J., Panda, D.K.: High performance virtual machine migration with RDMA over modern interconnects. In: Proceedings of the IEEE International Conference on Cluster Computing (Cluster'07), pp. 11–20 (2007)
15. Lange, J.R., Dinda, P.A.: Transparent network services via a virtual traffic layer for virtual machines. In: Proceedings of the 16th International Symposium on High Performance Distributed Computing, pp. 23–32 (2007)
16. Clark, B., Deshane, T., Dow, E.M., Evanchik, S., Finlayson, M., Herne, J., Matthews, J.N.: Xen and the art of repeated research. In: USENIX Annual Technical Conference, FREENIX Track, pp. 135–144 (2004)
17. Kivity, A., Kamay, Y., Laor, D., Lublin, U., Liguori, A.: KVM: the Linux virtual machine monitor. In: Proceeding of the 2007 Ottawa Linux Symposium 2007, pp. 225–230 (2009). Container-based virtualization for linux. [Online]. Available: https://www.openvz.com/

18. Bose, S.K., Brock, S., Skeoch, R., Rao, S.: CloudSpider: Combining replication with scheduling for optimizing live migration of virtual machines across wide area networks. In: 2011 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 13–22. IEEE (2011)
19. Travostino, F., Daspit, P., Gommans, L., Jog, C., De Laat, C., Mambretti, J., Monga, I., Van Oudenaarde, B., Raghunath, S., Wang, P.Y.: Seamless live migration of virtual machines over the MAN/WAN. Future Gener. Comput. Syst. **22**(8), 901–907 (2006)
20. Huang, W., Gao, Q., Liu, J., Panda, D.K.: High performance virtual machine migration with RDMA over modern interconnects. In: 2007 IEEE International Conference on Cluster Computing, pp. 11–20. IEEE (2007)
21. Bradford, R., Kotsovinos, E., Feldmann, A., Schiöberg, H.: Live wide-area migration of virtual machines including local persistent state. In: Proceedings of the 3rd International Conference on Virtual Execution Environments, pp. 169–179. ACM Press (2007)
22. Luo, Y., Zhang, B., Wang, X., Wang, Z., Sun, Y.: Live and incremental whole-system migration of virtual machines using block-bitmap. In: Proceedings of Cluster 2008: IEEE International Conference on Cluster Computing. IEEE Computer Society (2008)
23. Sapuntzakis, C.P., Chandra, R., Pfaff, B., Chow, J., Lam, M.S., Rosenblum, M.: Optimizing the migration of virtual computers. ACMSIGOPS Oper. Syst. Rev. **36**(S1), 377–390 (2002)
24. Sivathanu, S., Liu, L., Yiduo, M., Pu, X.: Storage management in virtualized cloud environment. In: 2010 IEEE 3rd International Conference on Cloud Computing, pp. 204–211. IEEE (2010)
25. van der Velde, O.A., et al.: Spatial and temporal evolution of horizontally extensive lightning discharges associated with sprite-producing positive cloud-to-ground flashes in northeastern Spain. J. Geophys. Res.: Space Phys. **115**(A9), (2010)
26. Zhang, X., Chen, L., Tong, Y., Wang, M.: EAGRE: towards scalable I/O efficient SPARQL query evaluation on the cloud. In: 2013 IEEE 29th International Conference on Data Engineering (ICDE), pp. 565–576. IEEE (2013)
27. Okafor, K.C.: Dynamic reliability modeling of cyber-physical edge computing network. Int. J. Comput. Appl. 1–11 (2019)
28. Frénot, S., Ponge, J.: LogOS: an automatic logging framework for service-oriented architectures. In: 2012 38th Euromicro Conference on Software Engineering and Advanced Applications, pp. 224–227. IEEE (2012)
29. Yu, X., Joshi, P., Jianwu, Xu., Jin, G., Zhang, H., Jiang, G.: Cloudseer: workflow monitoring of cloud infrastructures via interleaved logs. ACM SIGARCH Comput. Archit. News **44**(2), 489–502 (2016)
30. Ghoshal, D., Canon, R.S., Ramakrishnan, L.: I/o performance of virtualized cloud environments. In: Proceedings of the Second International Workshop on Data Intensive Computing in the Clouds, pp. 71–80 (2011)
31. Mao, M., Humphrey, M.: A performance study on the VM startup time in the cloud. In: 2012 IEEE Fifth International Conference on Cloud Computing, pp. 423–430. IEEE (2012)
32. Sobel, W., Subramanyam, S., Sucharitakul, A., Nguyen, J., Wong, H., Klepchukov, A., Patil, S., Fox, A., Patterson, D.: Cloudstone: multi-platform, multi-language benchmark and measurement tools for web 2.0. Proc. CCA **8**, 228 (2008)

Chapter 28
Real-Time Area-Based Traffic Density Calculation Using Image Processing for Smart Traffic System



S. Rakesh and Nagaratna P. Hegde

Abstract Traffic is one major problem in metropolitan cities and capital cities of all the states. As more people moving to the cities, the number of vehicles on the road is increasing rapidly. The traffic signals controlled by the traffic police manually have not proved to be efficient. Also the static predefined set time for traffic signals at all conditions whether the traffic density is heavy or low has not solved this problem. One solution for this problem is to develop a smart traffic management system which works by calculating the density of the vehicles on the road using real time video processing technique. We have used Mixture of Gaussian (MoG) algorithm for background subtraction technique and foreground detection, then generated capacity maps using which we found the density of traffic at that particular instant. The traffic lights at junctions are controlled dynamically based on the traffic density detected from the video feeds.

28.1 Introduction

The density of the vehicles when travelling strengthens slowly subsequently for the optimum use of existing street potential is rather notable to care for all the traffic flow appropriately. Traffic jams have wound up being sizable trouble specifically in the metro locations in addition to principal city regions. The primary cause is the fast increase in the populace of the metropolitan areas as well as principal city urban locations that subsequently boost vehicular travel, which makes traffic congestion

S. Rakesh
Department of Information Technology, Chaitanya Bharathi Institute of Technology (CBIT)(A), Gandipet, Hyderabad, Telangana 500075, India
e-mail: srakesh_it@cbit.ac.in

N. P. Hegde (✉)
Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_28

issue [1–5]. Because of these traffic jams, the transportation expense is boosting, the vacation possibility is actually rearing and also power intake is also enhancing [2]. Traffic additionally generates loads of different other censorious complications as well as worries which right effect private regular everyday lives and also sometimes trigger for the loss of life [6–8]. For instance, if there is a rescue when driving together with a severe individual on board. Because situation, if that rescue gets embedded notable traffic congestion at that point there are truly high options that the important individual may not arrive at the university hospital or hospital quickly. Therefore it is unbelievably vital to design along with making a wise traffic administration device which regulates the traffic intelligently to steer clear coming from wrecks, accidents and also traffic congestion [7, 8]. The main cause of traffic in developing nation is bad in addition to ineffective traffic keeping track of system which impacts the traffic flow. For example at a traffic light shared if one side road has a lot less traffic as well as the opposite service road along with huge traffic but the size of green light for every side is very same at that point this activates the rubbish of offered sources in addition to additionally this mishandles. Using thinking about the above case if the road or road along with higher traffic should activate the green stoplight for a longer time frame than the road or even lane together with very little traffic density at that point it decreases the options of traffic congestion adequately.

There are bunches of procedures prepared to design a smart and intelligent traffic management system, as an example, a fuzzy-based traffic controller in addition to morphological edge detection method is planned in [1]. This approach is based on the size of the density of the traffic by affiliating the recorded online traffic image with a recommendation fixed or even picked graphic. If the difference is a whole lot, even more, afterward greater traffic density is recognized. In [9], an additional method is suggested to create a smart traffic unit, which is based on the four-lane road system through which traffic light chance is allowed based upon a lot of vehicles when driving or maybe lane. The operation recommended in [10] is based upon the neural networks, which identifies the vehicles as well as likewise density of traffic through refining the snatched traffic video recordings. In [11], an added strategy is used which is based upon identifying the traffic capability by reviewing the two images, i.e., captured on the web traffic image the static reference photo.

This paper presents a traffic density analyzer task based on calculating the percentage of vehicles present in the recorded real-time traffic image, which provides our company with extra precise details for traffic light assortment creating.

The paper is taken care of as adhered to: Sect. 28.2 explains the system design, i.e., worrying MoG protocol, Sect. 28.3 details worrying the application procedure of the system, Sect. 28.4 explains concerning the results. Ultimately, Sect. 28.5 deals with involving conclusion as well as also future work abided by due to the important referrals made use of in this particular job.

28.2 System Model

Background subtraction is a mostly used moving object detection operation which diverse the foreground objects coming from the chosen frame about a static reference frame. The static reference frame is the background model, which is looked at as the previous frame.

$$|I(x,\ y,\ t) - I(x,\ y,\ t-1)| > \text{Th} \tag{28.1}$$

The threshold Th well worth is prepared manually. The difference at the facility of the absolute value of the decided frames needs to be lot larger than the threshold market value.

Depending on to [12], the algorithm is positioned in a variety of categories, coming from history-based protocols to adaptive learning procedures. Our team selects the versatile learning algorithm above history-based algorithm [13]. Flexible learning algorithm needs to have to possess simply a background model as well as also automatically updates the design based upon the present frame. As a result, it demands fairly a lot fewer memory data transactions to care for. On the contrary, history-based algorithm gathers all the video audio frames and also frequently accesses the frame history, subsequently asking for even more mind transmission capacity. Our group decided on a Mixture of Gaussian (MoG) algorithm, which is based on the adaptable learning algorithm which is featured in Fig. 28.1. To produce the background of a pixel, a Mixture of Gaussian (MoG) algorithm uses K Gaussian components along with mean, deviation in addition to weight. To locate changes in the brand-new frame, the new frame and likewise previous frames Gaussian components are compared to one another. The Gaussian components of the pixel are updated based upon the brand-new pixels learning factor. If all the Gaussian components carry out undoubtedly not match the new pixel values, at that point the all-new pixel is taken note of as the foreground.

28.3 Implementation

We gather a cache of 500 frames from the selected input video and use these frames to train our background model using openCV fixed function cv2.createBackgroundSubtractorMOG2(), which is based on the adaptive learning algorithm. It uses a technique to model each and every background pixel by a mixture of a suitable number of Gaussian distributions per pixel. The weights of these mixtures represent the duration of each individual pixel lodge on the screen. The pixels that lodge on the screen more time are more potential in being part of background.

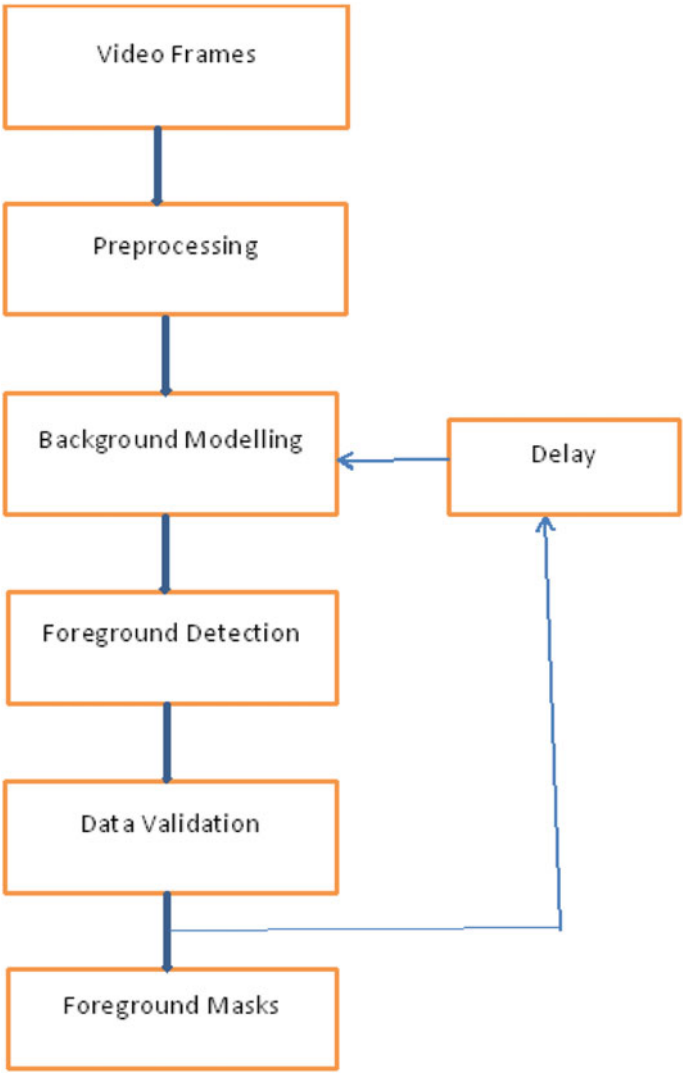


Fig. 28.1 Background subtraction for obtaining foreground mask

By the creation of a background model, we spot the foreground on the every next frame from the selected video input. The foreground still consists of noise, so we use filters on it to decrease the noise and make the foreground further useful for future use. The filters are closing, opening, erosion and dilation. Then we find the density of the traffic with the help of capacity maps which are shown in results section.

We have also send these traffic density values to an excel sheet shown in results section. Instead if we send these values to a smart traffic light management system,

then we can use the traffic lights effectively, i.e., to which line the density is more that side the duration of green light is more.

28.4 Results

We have tested our model by giving a selected traffic video as input and processed the video to generate outputs in the form of capacity maps which are shown in Figs. 28.2, 28.3, 28.4 and 28.5. The values displayed in these figures are traffic density values of frames generated from the input traffic video.

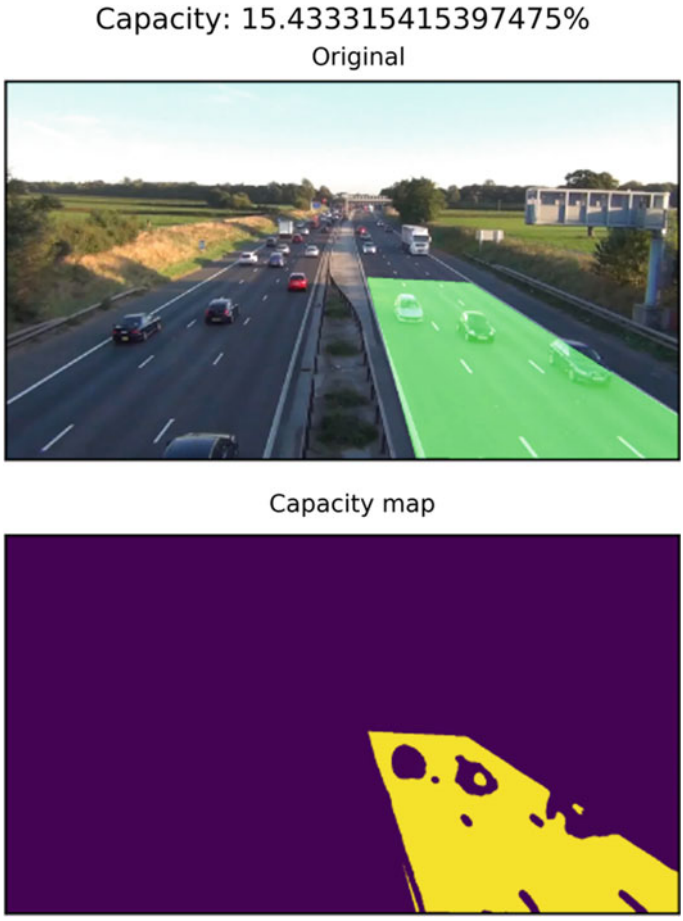


Fig. 28.2 Capacity map1

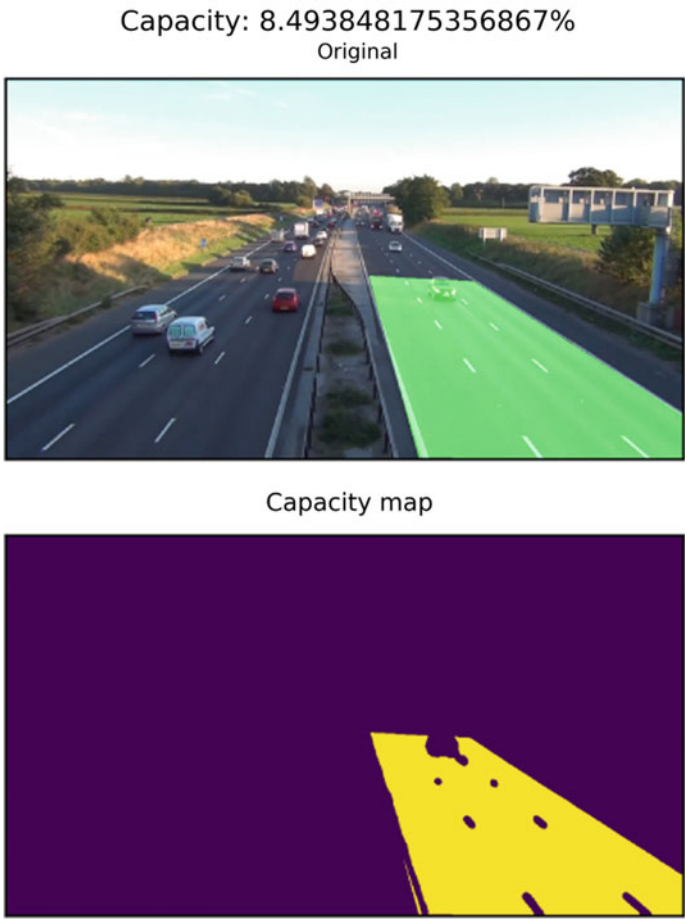


Fig. 28.3 Capacity map2

We have also send these traffic density values into an excel sheet which is shown in Fig. 28.6. We can change the time intervals of the frames according to our requirement and generate the capacity maps. These values are useful in taking decisions at the smart traffic control system, i.e., in general if the density of the vehicles is more in one side of the traffic signal than that road green traffic light should be more.

28.5 Conclusion and Future Scope

This paper covers a procedure for computing the density of traffic on the road by utilizing picture handling algorithm. Our team has used a Mixture of Gaussian (MoG)

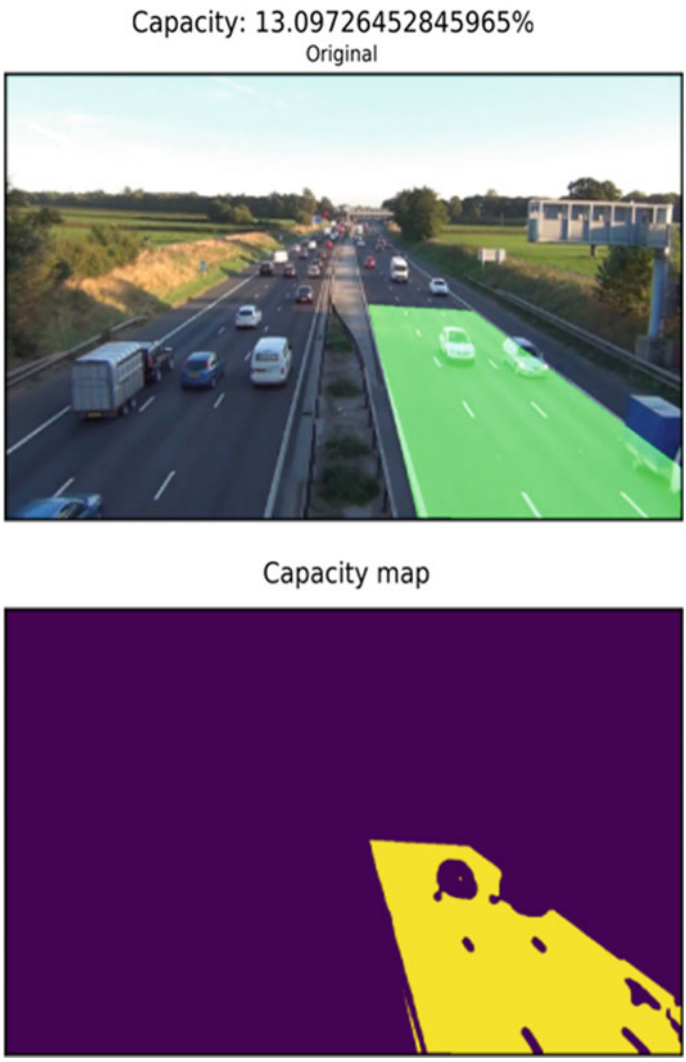


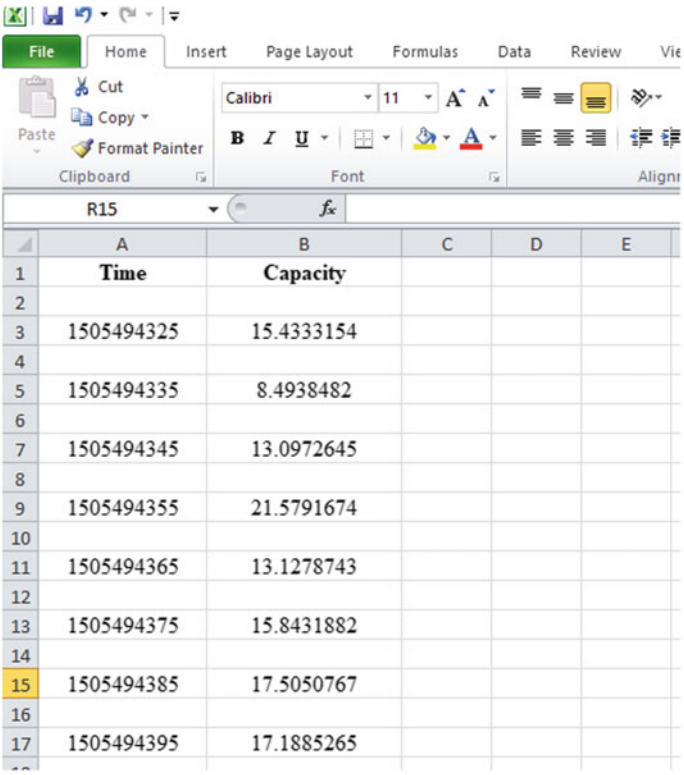
Fig. 28.4 Capacity map3

algorithm and also get the outcomes. The conveniences of the proposal procedure are that there is no demand to use extra complex sensing unit-based units such as RFIDs. The recommended device is inexpensive as it does not need to have an installation of some other sensing units or even units.

This work can be enhanced by sending these traffic density values to smart traffic management system and controls the signals accordingly.



Fig. 28.5 Capacity map4



	A	B	C	D	E
1	**Time**	**Capacity**			
2					
3	1505494325	15.4333154			
4					
5	1505494335	8.4938482			
6					
7	1505494345	13.0972645			
8					
9	1505494355	21.5791674			
10					
11	1505494365	13.1278743			
12					
13	1505494375	15.8431882			
14					
15	1505494385	17.5050767			
16					
17	1505494395	17.1885265			

Fig. 28.6 Traffic density values

References

1. Arora M., Banga, V.K.: Real time traffic light control system. In: 2nd International Conference on Electrical, Electronics and Civil Engineering (ICEECE'2012), pp. 172–176. Singapore (April 28–29, 2012)
2. Kanojia, S.B.: Real–time Traffic light control and Congestion avoidance system. Int. J. Eng. Res. Appl. (IJERA) **2**(2), 925–929 (2012)
3. Tayyab, M.: Implementation of restoration path using AODV in VANETs. Master's Dissertation at Brunel University London, UK
4. Venables, A.J.: Evaluating urban transport improvements. J. Transp. Econ. Policy **41**(2), 173–188 (2007)
5. Gärling, T., Schuitema, G.: Travel demand management targeting reduced private car use. J. Soc. Issues **63**(1), 139–153 (2007)
6. Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A.: Review of road traffic control strategies. Proceedings IEEE **91**(12), 2043–2067 (2004)
7. Vigos, G., Papageorgiou, M., Wang, Y.: Real-time estimation of vehicle-count within signalized links. J. Transp. Res. Part C: Emerg. Technol. **16**(1), 18–35 (2008)
8. Szeto, M.W., Gazis, D.C.: Application of Kalman Filtering to the Surveillance and Control of Traffic Systems. J. Transp. Sci. **6**, 4419–4439 (1972)
9. Dangi, V., Parab, A., Pawar, K., Rathod, S.S.: Image processing based intelligent traffic controller. Undergraduate Acad. Res. J. (UARJ) **1**(1), (2012)

10. Ozkurt, C., Camci, F.: Automatic traffic density estimation and vehicle classification for traffic survillance systems using neural networks. J. Math. Comput. Appl. **14**(3), 187–196 (2009)
11. Gupta, P., Purohit, G.N., Pandey, S.: Traffic load computation for real time traffic signal control. Int. J. Eng. Adv. Technol. 2(4), (2013)
12. Tabkhi, H., Bushey, R., Schirner, G.: Algorithm and architecture co-design of Mixture of Gaussian (MoG) background subtraction for embedded vision. In: 2013 Asilomar Conference on Signals, Systems and Computers, pp. 1815–1820. Pacific Grove, CA (2013)
13. Elgammal, A.M., Harwood, D., Davis, L.S.: Non-parametric model for background subtraction. In: Proceedings of the 6th European Conference on Computer Vision-Part II, pp. 751–767 (2000)

Chapter 29
An Association Mining Rules Implemented in Data Mining



B. Varija and Nagaratna P. Hegde

Abstract Data mining has created a great impact on the business organizations in different ways. The main purpose of mining is nothing but extracting useful and interesting knowledge-based patterns from large amount of data or information which is present in the data warehouses. The mining is nothing but identifying the frequent itemsets purchased by a customer from the store and by comparing the relationship between the itemsets that the analysis is been done. In order to identity the relationship, we introduce different association rule-based algorithms. We concentrate more on promising studies about the latest trends and technologies which are been implemented in the association mining rule. In this paper, we focus on association rule by using different algorithm to justify the rules.

29.1 Introduction

Data mining comprises of many techniques which are emerging day by day in each and every field either it may be the education, health, pharmacy, business, etc. Mining has changed the perspective of thinking of human being because data can be dumped from any source in the world sitting at one particular place, the data can be integrated from multiple sources, and that data can be placed at particular place which is nothing but our data warehouse. Online analytical processing can be used in the data mining which is nothing but analysis is been done with the help of the historical data which is been stored in the different repositories.

The main techniques of data mining are discussed as the association rule, clustering, classification, and outlier analysis and anomaly detection. Here purely, we would like to focus on the association mining rule which plays a vital role in data mining because of this mining rule the business organizations are able to identify

B. Varija
Sri Indu College of Engineering and Technology, Ibrahimpatnam, Hyderabad, Telangana, India

N. P. Hegde (✉)
Vasavi College of Engineering, Ibrahimbagh, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_29

the customer requirement in order to get benefit in their business. Association rules concentrate on the main perspective which is known as "market basket analysis", because it is considered as the one of the basic concept of the mining which will be further discussed in the paper.

29.2 The Association Rules

To understand the association rule, I would like to focus on the simple example, i.e., if we take into consideration a departmental store where we can find all household thinks so when a person purchase an item, the owner of the store will check the relationship between the itemsets purchased by the individual customers in order to draw what are the combination of products purchased by the different customers when they enter into the store. An association rule relays on two thinks; they are antecedent and consequent.

If customer buys beer he is likely to buy diaper

If we observe the above sentence, easily we can understand that here beer is the antecedent and diaper is consequent. The above example is given by the thorough analysis of the datasets by the company owner to improve the customer service and also improve the revenue of the company. To analyze the relationship between the itemsets, two parameters are taken into consideration, and they are observed.

$$\text{Support (A)} = \frac{\text{Number of transaction in which A appears}}{\text{Total number of transactions}}$$

$$\text{Confidence(A} \rightarrow \text{B)} = \frac{\text{Support (A} \cup \text{B)}}{\text{Support (A)}}$$

29.3 Apriori Algorithm

The Apriori algorithm is used to fine the frequent itemsets by following certain sequence of steps, and to perform this technique, I have used two basic steps, and they are join and purne. The join step is used to generate $K + 1$ itemsets, whereas the purne step is used for scanning the database. The good advantage of using this basic technique is to easily understand the algorithm. This algorithm is the level-wise generation of the frequent itemsets. The main feature is anti-monotonocity of the support. We do have certain limitation in the apriori algorithm that is it is very slow. Here, we do have itemsets frequent itemsets, frequent pattern mining, and the main things that we have to remember which are confidence and support.

- Mining frequent patterns are those patterns that occur frequently in the sales transactions of the super market.
- Frequent itemsets are those that frequently appear together.
- Frequent subsequence patterns are those which occur frequently like when you purchase phone will purchase SIM card.
- Association analysis process involves uncovering the relationship between data and deciding the rules for performing associations between objects
- Correlation analysis is a mathematical technique that can show together and how strongly the pairs of attributes are related
- Pattern evaluation is defined as identifying increasing patterns which represent the knowledge.

29.4 Multi-level Associative Rule

Here, we discuss about the rules at high concept level, and the support increases from specialized to general method. Support decreases from general to specialize method. Confidence is not affected for general or specialized.

Using uniform support for all levels, there is only one minimum support t so no need to examine itemsets. If support is too high which implies Miss Low level associations which is shown in the below figure which specifies the minimum support when it is too low (Figs. 29.1 and 29.2).

- A common form of background knowledge that an attribute may be generated or specialized according to a hierarchy of concept.
- Rules which contain associations with hierarchy of concepts are called multi-level association rule.

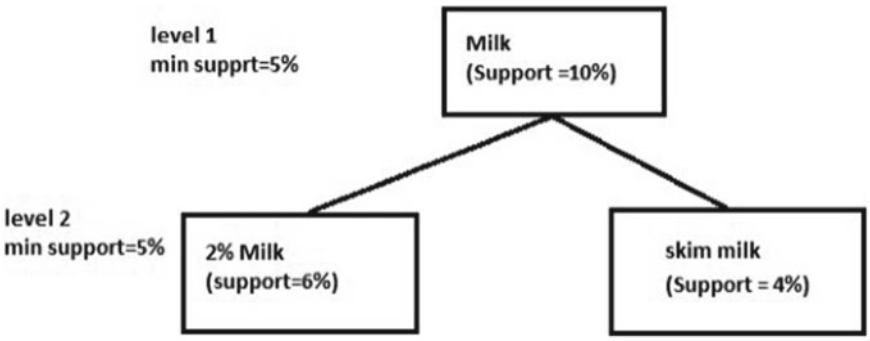


Fig. 29.1 Example of uniform minimum support for all levels

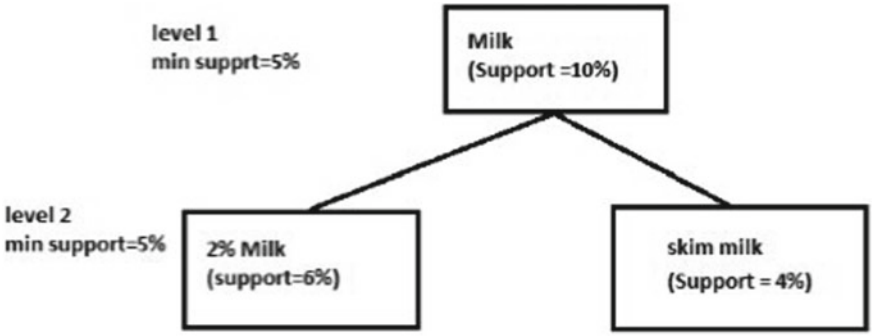


Fig. 29.2 Example of reduced minimum support for lower level

29.5 Multidimensional Association Rule

The multidimensional rule explains about the Boolean algebra. A Boolean matrix-based approach had employed to discover the frequent datasets. Apriori algorithm is used to prune the items. There are rules involving more than one dimension or predicates who try to buy products from the store. Attributes can be differentiated as the categorical or quantitative. Quantitative attributes are numeric and incorporates the concept hierarchy where the numeric attributes are discredited. We consider three different approaches of multidimensional association rules, and they are categorized as the following.

29.5.1 Mining Using Dynamic Discretization

The dynamic discretization is also known as quantitative where the numeric attributes are dynamically discredited (2D quantitative association mining rule). If we consider a simple example which is of age (X,"20…30) ˆ income (X,"30 K…60 K'0) → buys (X, "Laptop Computer").

29.5.2 Distance-Based Association Rule

This distance-based association rule consists of two steps involved in mining process, and they are as follows: Firstly, it performs clustering for finding the interval of attribute, and secondly, it obtains association mining rules by searching for the group of clusters that occur together.

29.6 Fuzzy Association Rule

The fuzzy rules are used to convert numerical attributes to fuzzy attributes; fuzzy apriori is quite slow algorithm which is available today. In data mining, we can get the best result using the fuzzy sets; the fuzzy rules are consider as the useful tools. The fuzzy rules are based upon fuzzy taxonomies. It is a computational model which is in the mathematical format. It was designed by L.zadeh. The fuzzy concept is an extent applicable in a situation; they are used in the traffic light, and the fuzzy technique is more particular used in the traffic control system [8].

29.7 Hybrid Dimensional Association Rules

To detect the fraud activities, we use the data mining. To perform the detection, hybrid method can be used. By using this method, it has less false discovery but provides with high accuracy. The rough set model is used for discovering the hybrid association rules. We have hybrid algorithm which are resource efficient and provide better performance.

29.8 Weighted Association Rule Mining

The WARM provides notion or weight to individual items that are solely on an item supported where this weighted rule is scalable and efficient. The weights can be arranged in the subjective manner. Linear model can be used to inculcate the weights and not only the linear model here we can also use the valency model. In the valency mode, first we consider the non-neighbor items, and the Gaussian elimination method is used. Secondly, it is been purified, and lastly, new model is developed.

29.9 Association Rule Mining in Different Databases

Firstly, we will talk about the association rule mining in large database where it is considered as the basic algorithm which can be performed on the sales transactions. To perform the association, here we use the very basic technique which is used in the apriori algorithm which is named as pruning. The best application we can discuss is sales transactions which can be performed in any shopping mall or super marker.

Secondly, its association rule mining in spatial database, this pattern can be collected from geospatial database. The spatial data consist of objects which are

spatial scene. Basic operations can be done to achieve the good result. The applications of spatial database are that it has the efficient storage system and prevents from the inconsistencies.

Thirdly, its association rule mining in distributed database is here to mine the data which we require large amount of data, so that mining can be done efficiently. The best way to solve problem in data mining is the distributed database mining. For proper mining, many distributed algorithms are used, and the different algorithms which are used for association rule mining in distributed database are:

Hierarchical
Collective Bayesian
Distributed association mining
Collective decision tree learning

Here, we have some applications of association rule mining techniques, and they are following.

Market basket analysis
Protein sequence
Medical diagnosis
Census data
CRM credit card

Not only this type of sectors but association mining plays vital role in different businesses by helping the business people in different ways, and here are few.

Helps business people to build good sales strategies
Helps business people to find the healthy marketing techniques
Self-life planning
In store organization

29.10 Conclusion

In this survey, we have got clear picture on different association mining rules. Here in order to overcome the limitations of the traditional association rule, we have discussed the new association rules, so that an efficient mining can take place in the business organizations which are working with the mining rules because in today's world the business is growing larger and the people want the easy way to develop their business by knowing the requirement of the users; for that reasons, the researchers are more concentrated to upgrade the techniques which are in the association mining rule. By using the emerging techniques of mining, we can easily solve the problems which are faced by the people in the business world, so this brings them the relief on their efforts.

References

1. Piatetsky-Shapiro, G.: Discovery, analysis, and presentation of strong rules. In: Piatetsky-Shapiro, G., Frawley, W.J. (eds.) Knowledge Discovery in Databases. AAAI/MIT Press, Cambridge, MA (2019)
2. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data—SIGMOD '93 (1993), p. 207. CiteSeerX: 10.1.1.40.6984. https://doi.org/10.1145/170035.170072. ISBN: 978-0897915922
3. Hahsler, M.: Introduction to a rules—a computational environment for mining association rules and frequent item sets. J. Stat. Softw. (2005)
4. Hahsler, M.: A Probabilistic Comparison of Commonly Used Interest Measures for Association Rules (2015). https://michael.hahsler.net/research/association_rules/measures.html
5. Hipp, J., Güntzer, U., Nakhaeizadeh, G.: Algorithms for association rule mining—a general survey and comparison. ACM SIGKDD Explor. Newsl. **2**, 58–64 (2000). CiteSeerX 10.1.1.38.5305. https://doi.org/10.1145/360402.360421
6. Brin, S., Motwani, R., Ullman, J.D., Tsur, S.: Dynamic itemset counting and implication rules for market basket data. In: Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data—SIGMOD '97 (1997) pp. 255–264. CiteSeerX 10.1.1.41.6476. https://doi.org/10.1145/253260.253325. ISBN 978–0897919111.
7. Omiecinski, E.R.: Alternative interest measures for mining associations in databases. IEEE Trans. Knowl. Data Eng. **15**, 57–69 (2003). CiteSeerX 10.1.1.329.5344. https://doi.org/10.1109/TKDE.2003.1161582
8. Aggarwal, C.C.; Yu, P.S.: A new framework for itemset generation. In: Proceedings of the seventeenth ACM SIGACT-SIGMOD- SIGART symposium on Principles of database systems—PODS '98, pp. 18–24 (1998). CiteSeerX 10.1.1.24.714. doi:https://doi.org/10.1145/275487.275490. ISBN: 978-0897919968.
9. Piatetsky-Shapiro, G.: Discovery, analysis, and presentation of strong rules. In: Knowledge Discovery in Databases, pp. 229–248 (1991)
10. Tan, P.-N., Kumar, V., Srivastava, J.: Selecting the right objective measure for association analysis. Inf. Syst. **29**(4), 293–313. CiteSeerX 10.1.1.331.4740. https://doi.org/10.1016/S0306-4379(03)00072-3
11. Tan, P.-N., Michael, S., Kumar, V.: Chapter 6: Association analysis: basic concepts and algorithms. In: Introduction to Data Mining. Addison-Wesley. ISBN 978-0-321-32136-7
12. Pei, J., Han, J., Lakshmanan, L.V.S.: Mining frequent itemsets with convertible constraints. In: Proceedings 17th International Conference on Data Engineering, pp. 433–442 (2001). CiteSeerX 10.1.1.205.2150. https://doi.org/10.1109/ICDE.2001.914856. ISBN: 978-0-7695-1001-9
13. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large databases Archived 2015–02–25 at the Wayback Machine. In: Bocca, J.B., Jarke, M., Zaniolo, C. (eds.) Proceedings of the 20th International Conference on Very Large Data Bases (VLDB), pp. 487–499. Santiago, Chile (September 1994)
14. Zaki, M.J.: Scalable algorithms for association mining. IEEE Trans. Knowl. Data Eng. **12**(3), 372–390 (2000). CiteSeerX 10.1.1.79.9448. https://doi.org/10.1109/69.846291
15. Zaki, M.J., Parthasarathy, S., Ogihara, M., Li, W.: Parallel algorithms for discovery of association rules. Data Min. Knowl. Disc. **1**(4), 343–373 (1997). https://doi.org/10.1023/A:1009773317876
16. Han: Mining frequent patterns without candidate generation. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. SIGMOD '00, pp. 1–12 (2000). CiteSeerX 10.1.1.40.4436. https://doi.org/10.1145/342009.335372. ISBN: 978-1581132175
17. Witten, Frank, Hall: Data Mining Practical Machine Learning Tools and Techniques, 3rd edn
18. Hájek, P., Havránek, T.: Mechanizing Hypothesis Formation: Mathematical Foundations for a General Theory. Springer-Verlag (1978). ISBN 978-3-540-08738-0
19. Webb, G.I.: OPUS: an efficient admissible algorithm for unordered search. J. Artif. Intell. Res. **3**, 431–465 (1995). (AAAAI Press, Menlo Park, CA online access)

Chapter 30
Wi-Fi Networking: A Novel Method for Stabilizing Internet Connectivity by Optimizing Wi-Fi Router Configuration



Vijay A. Kanade

Abstract Internet is playing a critical role in today's world. It has kept people connected regardless of the physical barriers. Business connections, communications, and remote work seem a possibility today due to the Internet. However, with the growth in smart devices and an exponential increase in the number of Internet users, there has been significant growth in the global Internet traffic which has exposed the limitations of the RF signals used in wireless communication in households, offices, and public places. In addition, these RF signals encounter disturbances as they travel through the transmission medium, i.e., air, and undergo attenuation. As a result, the RF signals received at the client end device seem to display an inconsistent and unstable flow, leading to an unsteady Internet connection for its users. The research paper discloses a novel solution to address the problem of inconsistent and unstable Internet connection observed in a Wi-Fi networking environment, wherein Wi-Fi routers are at play in a defined geography.

30.1 Introduction

Internet had become an indispensable part of our lives today. It has bettered the quality of human life by fueling technological developments with regard to communication, IoT, and many more fields. As per the July 2020 stats, around 4.57 billion users are active on the Internet [1]. With the advent and advancement in smart devices, it can only be expected that a greater number of users and devices may come online in the near future.

As Internet has reached the households, the corporate world has seen a cultural shift. Corporate companies are encouraging remote work as employers are benefitting from it. As per the recent stats, workplaces are embracing remote work alternatives more than ever before [2].

V. A. Kanade (✉)
Intellectual Property Research, Pune, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_30

Currently, the world is grappled with global pandemic of COVID-19, and such technological adoption is acting as a boon for the remote workforce.

The work-from-home environment is here to stay. However, such work culture change is slated to put pressure on the Internet connectivity paradigm. As employees stay home bound, Internet or Wi-Fi sharing with spouse, partner, roommate, or kids may become a norm. In such scenarios, the risk of handling lousy Internet connection looms large. With buffering video conference calls and delayed audio, the employees are bound to suffer. Further, movies and video games may take ages to download. In the worst cases, the connection may drop altogether or seem to slow down due to heavy bottleneck on Internet speeds, leading to unstable Internet connectivity.

The connection drop or slowness issue can be attributed to various reasons, such as problem with router or modem, problem with Wi-Fi signal, signal strength, or number of devices on a network that hog the bandwidth.

However, even if these problems are considered and provided with a solution, still the Internet connection may not stabilize. In addition, broadband router configuration errors, wireless interference, or several other technical problems can adversely affect the speed and stability of the Internet connection.

All these connectivity problems are essentially associated with the radio wave property of the wireless connectivity signal. Therefore, there seems to be a long-standing need to provide a technical solution to the prevailing radio wave problem posed by the Wi-Fi broadcasting medium, i.e., router.

30.2 Wi-Fi Networking

Wi-Fi networking technology allows devices to wirelessly connect to the Internet. The wireless connection facilitates data exchange among devices distributed across the network. Wi-Fi is based on IEEE 802.11 family of standards that specify the set of protocols for implementing wireless communication in a network at various frequencies.

There have been various Wi-Fi versions, with the first version being released in 1997 and updated in 1999. The latest Wi-Fi version, called Wi-Fi 6 or 802.11ax, claims to achieve speeds of 11 Gbit/s (i.e., theoretical data rates).

The Wi-Fi-enabled devices exchange information by sending and receiving radio waves. These radio waves define a type of electromagnetic radiation in the electromagnetic spectrum, having longer wavelength and lesser frequency than microwaves and infrared waves. Specifically, these radio waves have frequencies ranging between 2.4 and 5 GHz [3].

Wi-Fi network consists of a central device called as router that provides Internet access to the connected devices within the network. Diverse devices such as smartphones, PCs, and various Wi-Fi-enabled gadgets or appliances (e.g., Jio-Fi device) can act as router, as long as the connected devices are within its range or network.

While operating on 2.4 GHz frequency, Wi-Fi routers can reach up to 150 feet (46 m) indoors and 300 feet (92 m) outdoors. Routers broadcasting at 5 GHz

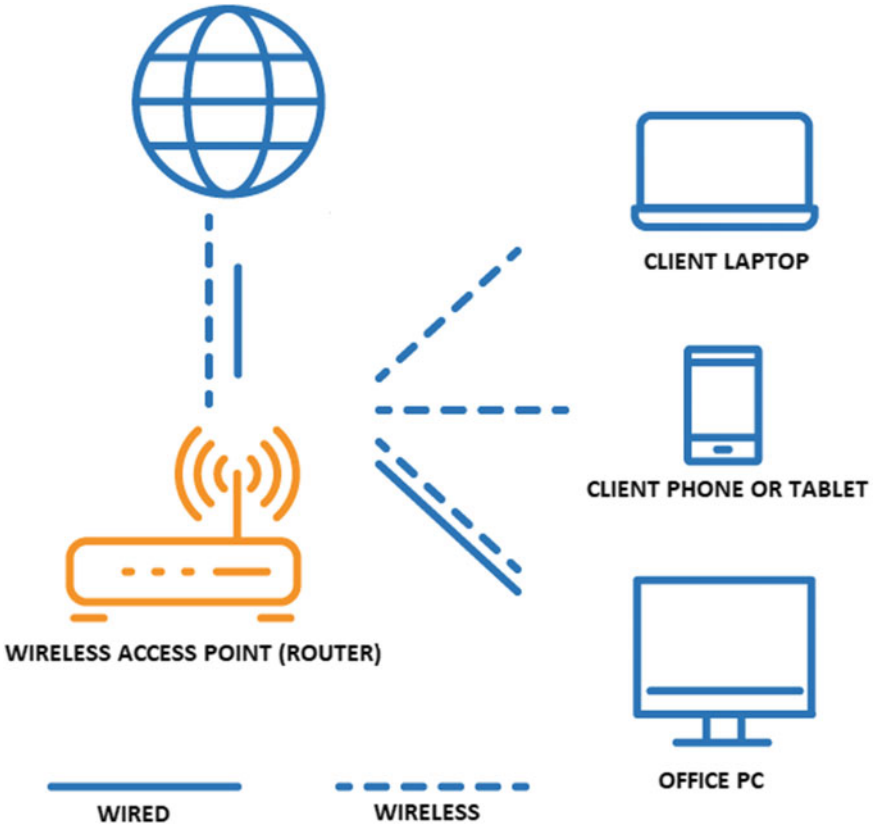


Fig. 30.1 Wi-Fi networking [4]

frequency have smaller range coverage but are generally immune to interference caused by other connected devices. Wi-Fi networking is illustrated in Fig. 30.1.

30.3 Radio Waves

A radio wave is an electromagnetic signal that carries information through the air over long distances. These radio waves are termed as radio frequency (RF) signals. RF signals travel through the air (as like waves in the ocean) at a high frequency. These signals carry data such as music, FM radios, and television videos over the air. A simple radio wave is represented in Fig. 30.2.

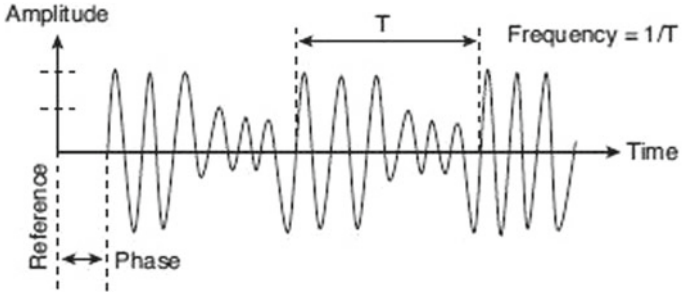


Fig. 30.2 Radio wave having amplitude, frequency, and phase elements [5]

30.3.1 RF System Components

RF system components enable the propagation of radio waves over the medium, i.e., air. A client device is equipped with a transceiver and an antenna for data transmission and reception. The transmission medium may include obstacles, such as walls, ceilings, or objects such as furniture and chairs. The standard RF system is disclosed in Fig. 30.3.

30.3.1.1 RF Transceiver

RF transceiver comprises a transmitter and a receiver. The transmitter transmits the radio waves from the source, and the receiver receives the waves from the destination. RF transceiver is an integral part of any client device that participates in the data exchange with the other devices.

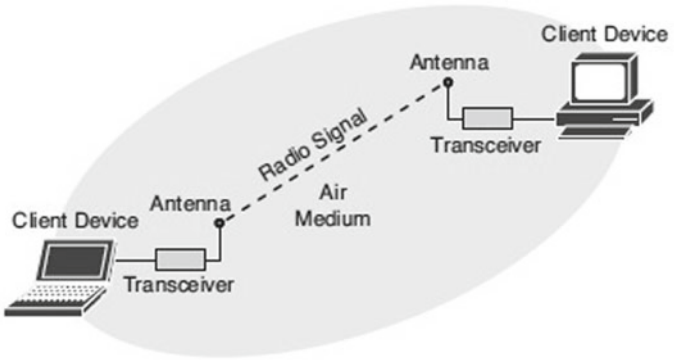


Fig. 30.3 RF system components—RF transceiver, antenna, and a transmission medium [5]

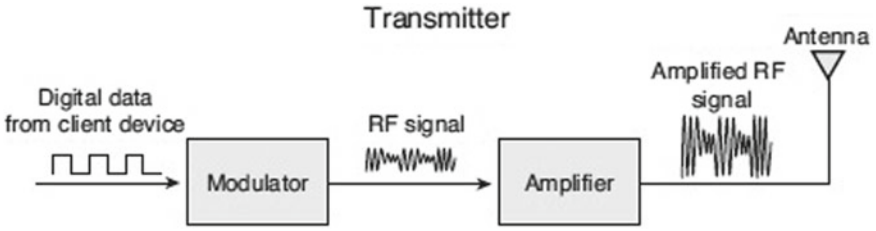


Fig. 30.4 A transmitter having a modulator, an amplifier, and an antenna [5]

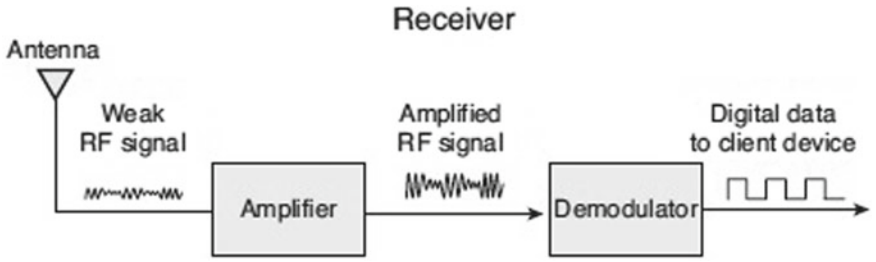


Fig. 30.5 A receiver having an antenna, an amplifier, and a demodulator [5]

The transmitter consists of a modulator and amplifier, wherein the modulator converts electric digital signal representing information (data bits, i.e., 1s and 0s) into radio wave of desired frequency that propagate through the air and the amplifier increases the amplitude of the very radio wave signal to a desired transmit power before it is fed to the antenna. A transmitter component of the transceiver is represented in Fig. 30.4.

A receiver at the destination client device detects the RF signal sent by the sender and demodulates it into data types that it is compatible with. The radio wave received at the receiver must compulsorily have amplitude that is above the receiver sensitivity, so that the receiver can interpret the signal, or decode it. Here, the minimum receiver sensitivity depends on the data rate of the RF signal (Fig. 30.5).

30.3.1.2 RF Modulation

RF modulation converts digital data, i.e., binary 1s and 0s, that represent a document or e-mail into RF signal and transmits it through the air. This includes conversion of digital data signal into analog signal. As modulation progresses, the digital data signal is superimposed onto a carrier signal, wherein the carrier frequency is the radio wave with a specific frequency. As a result, the digital data rides over the carrier. Further, the modulation signal varies the carrier signal in a manner that represents the data to be communicated. At the receiving device, the antenna couples the modulated signal into a demodulator, thereby deriving the data signal from the signal

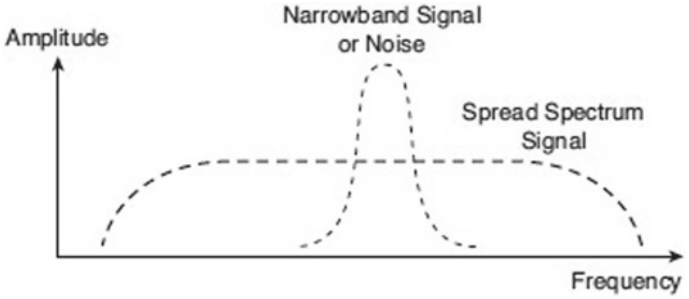


Fig. 30.6 Portion of RF spectrum occupied by spread spectrum [5]

carrier. Various modulation techniques include amplitude-shift keying, frequency-shift keying (FSK), phase-shift keying (PSK), and quadrature amplitude modulation (QAM).

30.3.1.3 Spread Spectrum

Some transceivers after modulating the digital signal into an analog signal spread the modulated carrier over a wider spectrum to reduce the possibility of outward RF signal and inward RF signal interference. Here, a signal's power is spread over a wide band of frequencies, as represented in Fig. 30.6.

30.4 Research Developments

A team of researchers from the University of Sydney's Australian Institute for Nanoscale Science and Technology have developed a chip-scale optical device that achieves radio frequency signal control at sub-nanosecond time scale. The research development potentially provides broader bandwidth to more users.

The developed device appears to be a critical part of the antenna system, wherein the device creates multiple configurable beams that can access more users in real-world environment. The device uses tunable delay lines on chip that reconfigure direction of the beam at nanosecond speed. This implies that each beam carrying digital data over the RF signal can be targeted to each user, thereby providing high bandwidth capacity at a given time [6].

The above research relates to reconfiguring the antenna system in order to provide better bandwidth capacity to each individual user accessing the Internet in a network. However, the research fails to address the grave problems such as Wi-Fi signal strength, router or modem technical issue, transmission medium issue, wherein the medium may include obstacles like walls or objects such as furniture. Thus, the

Internet connectivity issue may still suffer even if the reconfigured antenna system is used by the devices operating in the network.

30.5 Optimizing Wi-Fi Router Configuration

The research proposal discloses a change in the hardware configuration of the Wi-Fi routers that are used in households, offices, or personal spaces. The Wi-Fi router circuit is integrated with adjustable and tunable delay line that consist of one or more on-chip delay elements.

The adjustable delay line deployed within the router forms a critical component of clock synchronization circuits such as a delay locked loop (DLL). The delay line inserts a precise delay into the path of a signal, wherein the produced delay may be adjusted by an analog voltage or digital control word. The delay element(s) within the tunable RF delay line create a time difference between input and output signal. By tuning the delay line, this time difference can be continuously changed, at the order of a billionth of a second [7].

The DLL circuitry receives a reference clock signal (XCLK) at an input buffer, which is further provided to a delay line as a buffered clock signal (CLKIN). The delay line consists of one or more on-chip delay elements that induce delay in the RF signal passing through them. The entry point of the buffered clock signal CLKIN may be adjusted by harnessing a number of delay elements to provide a lock through a range of frequencies, input voltages, etc. The output of the delay line is connected to an output buffer, wherein the RF signal is broadcasted via this output buffer (Fig. 30.7).

Adjustable RF delay line containing one or more delay elements is selected based on the following specifications (Table 30.1).

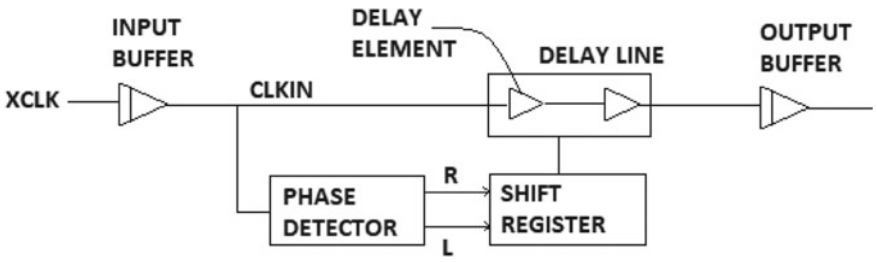


Fig. 30.7 DLL circuitry

Table 30.1 Specifications of the delay line

Sr. no.	Specifications	Details
1	Frequency	This is the frequency of interest at which a time delay is required. Generally, delay lines are optimized for a specific frequency range
2	Delay	This is the amount of delay that is provided by the delay line. Delays are usually in the micro- to nanosecond range
3	Insertion loss (dB)	This is the loss in amplitude of the signal after it is delayed by the delay line
4	Power level	This is the power level that the delay line can handle. It is important to choose a delay line that can handle a power level higher that what will be used in the RF system

30.5.1 Use Case

Consider a scenario, wherein the router receives wireless RF signal from the Internet service provider through the air. The received RF signal is allowed to pass through the delay line containing delay element(s). These delay elements buffer the received RF signal along its circuitry by inducing phase or time delays. As the received first RF signal is buffered in the delay element, the second RF signal and the consequent RF signals get queued within the delay elements. This queuing of RF signals in the on-chip delay elements of the delay line thereby facilitates storage of RF waves. Thus, the RF delay elements store radio waves by inducing time or phase delays for a brief period before broadcasting them. These queued RF signals are later broadcasted by the router for facilitating Internet connectivity in the nearby devices.

The diagrammatic representation of the proposed RF router 'A', leading to the buffered RF waves 'B' within the delay line, is as disclosed in Fig. 30.8.

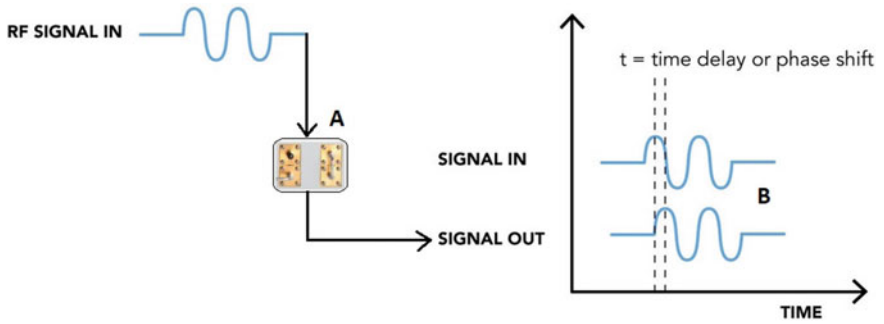


Fig. 30.8 Optimized router configuration

30.5.2 Algorithm

(a) **Step 1:** RF signal is received as input by the router.
(b) **Step 2:** RF delay line within the router induces time or phase delay (depending on the router configuration) in the signal.
(c) **Step 3:** RF input signal is buffered along the delay line containing the delay elements.
(d) **Step 4:** Buffered RF signal is broadcasted by the output buffer of the router.

30.6 Advantages

The optimized router configuration disclosed in the research has various advantages enlisted below:

1. Delay elements in the router provide stable Internet connection throughout the connectivity session due to buffered RF signal.
2. Internet connection may not drop frequently due to inconsistency on the RF signal transmission through the medium, i.e., air.
3. Increased data speeds with consistent supply of adjusted RF signal via delay line in the router.
4. The 'bandwidth bottleneck' has caused the demand for bandwidth to rapidly shoot up against the limited supply available today and is estimated to increase by 22% each year. The research work solves the problem of bandwidth bottleneck, as the RF signals are equally spaced and queued within the delay line and broadcasted at regular intervals.
5. Seamless and consistent Internet connection as the vulnerable radio wave property is adjusted by the tunable RF delay line, wherein the delay line acts as the RF storage device.

30.7 Conclusion

The research paper provides a novel solution for the problem of inconsistent and unstable Internet connection accessed by the users at households, offices, or public places via Wi-Fi routers. The proposed research provides a technical solution to the radio wave problem posed by the Wi-Fi broadcasting medium, wherein the Wi-Fi signals are broadcasted at regular intervals by integrating the delay line within the Wi-Fi router for consistently providing stable Internet connection to its users in the vicinity.

Acknowledgements I would like to extend my sincere gratitude to Dr. A. S. Kanade for his relentless support during my research work.

Conflict of Interest The authors declare that they have no conflict of interest.

Declaration We have taken permission from competent authorities to use the images/data as given in the paper. In case of any dispute in the future, we shall be wholly responsible.

References

1. Clement, J.: Worldwide Digital Population as of July 2020. (July 24, 2020)
2. Bump, P.: 40 Remote Work Stats to Know in 2020, (Mar 12, 2020)
3. How Does WiFi Work? NetSpot
4. How Does Waves Work? Waves Guest Wi-Fi Manager
5. Cisco Press, WiFi Networking: Radio Wave Basics, (January 11, 2017)
6. Photonics Breakthrough Tunes Wireless Communications, (April 19, 2017)
7. Gomm, T.J. et. al.: Synchronization device with delay line control circuit to control amount of delay added to input signal and tuning elements to receive signal form delay circuit. US7111185B2, 2003-12-23

Chapter 31
Periocular Segmentation Using K-Means Clustering Algorithm and Masking



V. Sandhya and Nagaratna P. Hegde

Abstract Biometrics are used widely to authenticate Humans to access various systems and devices in real-world applications. The Periocular region can be used to recognize an individual for identity management. In this paper, we propose the method using the K-Means algorithm to cluster. The image and then apply masking operations to extract features from the image.

31.1 Introduction

User validation is a unique challenge in the field of Biometrics. The different Biometric methods used to authenticate a person are a face, fingerprint, iris, palm print, Handwriting, and sclera. In all of these methods, each individual has to possess unique features. In the present day, Face recognition is the most prominent authentication mechanism for an identity management system. We try to extract unique features of the face. Periocular is the region surrounding the eye that includes the upper eyelid, lower eyelids, sclera, iris, periocular skin, and eyebrow. Challenges that are associated with periocular recognition are persons aging, illumination levels and alignment. To overcome these challenges, the accuracy of Periocular recognition serves as an important task. Figure 31.1 shows a typical periocular region from an open-source database set the UBIPr.

V. Sandhya (✉)
Gitam School of Technology, Hyderabad, India
e-mail: svinayak@gitam.edu

N. P. Hegde
Vasavi College of Engineering, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_31

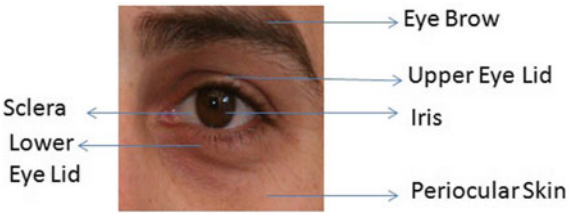


Fig. 31.1 Periocular image

31.2 Related Work

In [1] proposed OCLBCP method using dual-stream CNN to overcome limitations of the wild environment using the UBIPr database. In [2], the author presented an efficient way to detect and position the eye in the face using cascaded CNN. Deep Learning frameworks are used in [3] to extract complex features by applying spatial and color transformation used for data augmentation. The different features used are local features using various vital points and global features [4]. In [4], the authors present the databases used, such as UBIPr, FOCS, IMP: IITD Multispectral Periocular, CSIP, and also the different algorithms used for extracting the best regions for periocular recognition. To extract unique features from the face, LBP is used in [5], and to classify the distance city block is used in [6]. Image is segmented into clusters using an unsupervised K-Means clustering algorithm [7]. The subtractive clustering model is used to create data points. SIFT descriptors are applied to the image [8] to authenticate individuals using the periocular region in an unconstrained environment and GMM model. The four categories in periocular Biometrics are [9] classified into four categories, such as local, global, key-point, and shape appearance. Depending on the features needed to extract from any of the four classes can be applied. Periocular images can be captured in both NIR and the visible spectrum. As proposed in [10], the best regions of the periocular can be selected using the SFFS method. In [11] images, both the visible night images and near-infrared spectrum images are trained using two neural networks to implement a cross-spectral mechanism. Some of the challenging data sets are soft tissue identified in the periocular region. In [12], the proposed method is a hierarchical patch to identify soft tissues in the periocular region and even verified for twin data set. A data level fusion has been proposed in [13] to obtain periocular features eyebrow, eye, and skin regions. The periocular region has a rich texture and requires less user cooperation when compared to other biometric methods. The discriminate features for periocular recognition are proposed in [14]. The dimensions of LBP are reduced to half to transform into binary patterns in [15] an image retrieval system. An image can be divided into patches [16], and colors are matched to evaluate periocular points in the image. In [17], LBP and GIST are used to analyze the distance of images captured in periocular recognition.

31.3 Median Filtering

This low-pass filter is used to eliminate noise from an image that results in blurring of images. Determine the median of all pixels in the window and restore the central pixel with the median value. Kernel size is an odd number.

31.4 Canny Edge Detection Algorithm

Canny edge is a multi-stage edge detection algorithm.

1. Noise in the image can be removed by using 5 × 5 Gaussian Filters.
2. Sober Kernel is applied to the image to get horizontal gradients Gx direction and Vertical gradient direction Gy. Then we find the edge gradient and direction of each pixel.
3. We scan the image pixel by pixel to remove any pixel that does not form edges.
4. We try to determine edges and non-edges. The two thresholds are MIN and MAX**.** Edges with gradient values more than MAX are edged that appear, and edges with a value less than MIN are not edges and do not appear.

Given the periocular image (I) of size height (H) × width (W). The image I is clustered into K number of clusters. Suppose P (H, W) is a pixel and C_k denotes central cluster Pixels. After applying the K-Means algorithm, the resultant image is known as the clustered image (I').

31.5 K-Means Clustering Algorithm

1. Choose the number of clusters K to form for a given image.
2. Arbitrarily choose K points as central pixels from the image.
3. Compute distance for each pixel to the central pixels.
4. Allocate pixel to the nearest cluster central pixel with minimum distance when compared to all other cluster central pixels.
5. Determine the new cluster central pixel.
6. Repeat Step 3–5 to attain the desired number of iterations or expected accuracy.

31.6 Masking

To extract a cluster K of Periocular Image I, we apply masking operations on the clustered image and part of the image to be extracted. Pick the index label of the cluster (C_k) from the clustered image I'. Create a mask with cluster C_k central pixel coordinates as a circle. Obtain the coordinates of the selected cluster C_k of Image

I'. Label all the connected blobs of the cluster C_k mask. Create a mask for the entire cluster of blobs. Extract the cropped cluster Image of the Periocular Image *I*.

31.7 Proposed Method

In this method, we extract features of the periocular images such as eyebrow, periocular skin texture, iris, and sclera depending on the number of clusters *K* chose. Here $k = 7$, we extracted three features such as eyebrow, Iris, and upper eyelid and periocular skin of the Image I. Figure 31.2 shows the steps involved in extracting features from the given image *I*.

1. Periocular image (I) to be segmented is loaded, and apply image enhancement technique.
2. Select the number of Clusters K based on the number of features to be extracted from the Periocular image (I).
3. Pass the periocular Image I through Median filter to remove noise from the image.
4. Apply the canny edge detection algorithm to identify the edges of the periocular image.
5. Use the K-Means clustering algorithm to extract the features on the given periocular Image I and generate the clustered image I'.
6. Extract a specific feature of the clustered image by selecting the cluster C_k and apply mask operation.
7. Repeat step 7 to extract more than one feature of the Periocular Image I.

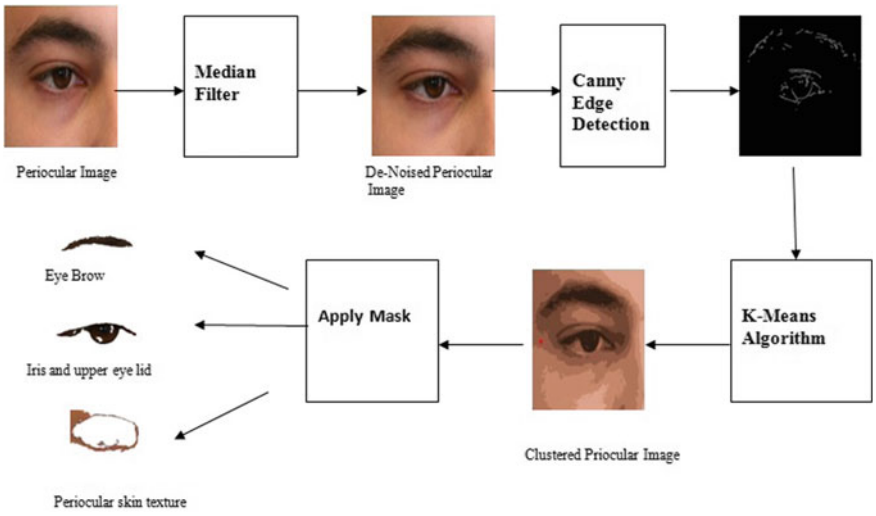


Fig. 31.2 Steps to extract periocular image features

31.8 Data Sets

See Fig. 31.3 [18, 19].

31.8.1 UBIRIS.V1

In this first version of the database, there are 1877 images gathered from 241 eyes in a less constrained environment.

31.8.2 UBIRIS.V2

In this database, there 11,000 images recorded at a distance in a mobile environment.

31.8.3 UBIPr

Is the updated data set of UBIRIS.v2? For periocular images. Images in the data set can be used specifically for periocular image processing and identification. The data set is open source.

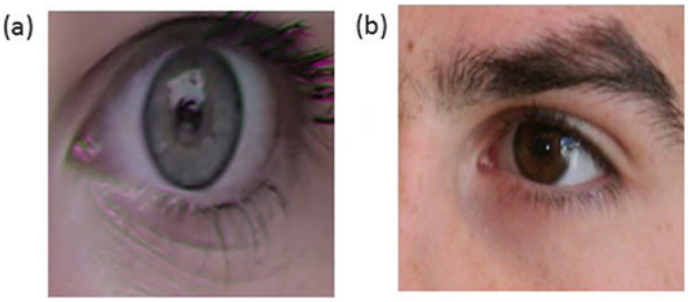


Fig. 31.3 Data sets **a** VSIOB samsung_day_light_short; **b** UBIPr

31.8.4 Visob

Dataset—The database contains images of eyes collected from 550 persons using Oppo, Samsung, and iPhone mobile devices. Images are captured in desperate environment conditions such as out-of-focus, eye-makeup, gaze directions, glasses, and discrete illuminations. In this paper, UBIPr and VISOB Samsung data set are used for implementation.

31.9 Results

We used periocular images from UBIPr and VSIOB Samsung_day_light_short. In the UBIPr data set, each image is of size 300×400, and in the VSIOB data set, images are of size 160×240. Images enhancement is done to attain images from different data sets of the same size. The resultant enhanced images are of size 400×400. To discard any noise in the image, we applied a Median filter, which results in smoothing of the image. Then canny edge detection algorithm is used to identify edges in the periocular image (I). we choose the number of clusters K as 9 to extract features from the Periocular Image such as eyebrows, Iris, upper eyelid, lower eyelid, periocular skin, and sclera. K-Means clustering algorithm is used to cluster the periocular image into different features. Depending on the intensity of the pixel values, clusters and central pixels are formed. We applied K-Means for ten iterations. To the clustered image I masking is applied to get the selected feature from the image. As shown in Fig. 31.4, two images from the VSIOB data set and one image from the UBIPr data set is chosen. We tried to extract different features from the three selected images of the data set. Figure 31.4a the upper eyelid of the image is extracted by selecting the central pixel of the cluster. Figure 31.4b eyebrow is extracted from the image. Figure 31.4c we extracted the iris feature from the image. To extract the features from the three images, we applied masking operation by selecting a central pixel from the corresponding cluster.

31.10 Conclusion

We applied K-Means clustering and masking to segment the periocular image. The resultant clustered image is used to extract the important features from the image. In this method, masking of the image is the crucial step in extracting features. Using the Features extracted, and we can create a data set for periocular images. The proposed method can serve as a base for applying different machine learning algorithms to solve image classification problems.

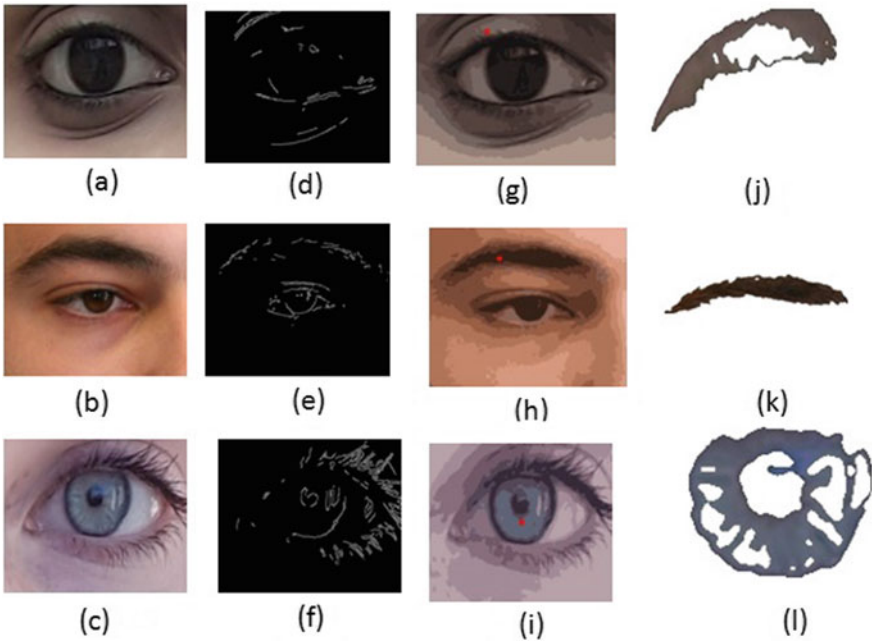


Fig. 31.4 Images **a**, **c** from VSIOB samsung_day_light_short; Image **b** from UBIPr data set; **d**–**f** Canny edges of the images; **g**–**i** K-Means clustered image; **j** upper eye lid of (**a**); **k** eye brow if (**b**); (**j**) iris of (**c**)

Acknowledgements We like to express gratitude to the University of Missouri-Kansas City for providing access to **the VISOB** data set and the Department of Computer Science, the University of Beira Interior, for providing access to the **UBIPr** dataset.

Declaration We, authors declare that we have taken required permissions from the respective person/authority to use the image/dataset in our work. We shall be responsible in case any dispute arises in the future with regard to this.

References

1. Ching, L., Tiong, O., Lee, Y., Beng, A., Teoh, J.: Periocular Recognition in the Wild: Implementation of RGB-OCLBCP Dual-Stream CNN, Applied Sciences, (July 2019)
2. Li, B., Fu, H.: Real-time eye detector with cascaded convolutional neural networks. Hindawi Appl. Comput. Intell. Soft Comput. (2018)
3. Proenc, H., Neves, J.C.: Deep-PRWIS: periocular recognition without the iris and sclera using deep learning frameworks. IEEE Trans. Inf. Forensics Secur. **13**(4), (April 2018)
4. Fernandez, F.A., Bigun, J.: A Survey on Periocular Biometrics Research. Elsevier (2018)
5. Kumar, K.K., Pavani, M.: LBP Based Biometric Identification using the Periocular Region. IEEE (2017)
6. Miller, P.E., Rawls, A.W., Pundlik, S.J.: Personal Identification Using Periocular Skin Texture. ACM (2010)

7. Dhanachandra, N., Manglem, K., Chanu, Y.J.: Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. In: Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)
8. Monteiro, J.C., Cardoso, J.S.: Periocular recognition under unconstrained settings with universal, background models. In: Proceedings of the International Conference Bio-inspired Systems and Signal Processing (BIOSIGNALS-2015)
9. Woodard, D.L.: Periocular-Based Biometrics**,** Encyclopedia of Biometrics (2014)
10. Alonso-Fernandez, F., Bigun, J.: Best Regions for Periocular recognition with NIR and Visible Images. IEEE (2014)
11. Sharma, A., Verma, S., Vatsa, M., Singh, R.: On Cross-Spectral Periocular Recognition. IEEE (2014)
12. Mahalingam, G., Ricanek, K., Jr.: LBP-based periocular recognition on challenging face datasets. EURASIP J. Image Video Process. (2013)
13. Moreno, J.C., Surya Prasath, V.B., Proenca, H.: Robust Periocular Recognition by Fusing Local to Holistic Sparse Representations. ACM (2013)
14. Ambika, D.R., Radhika, K.R., Seshachalam, D.: The Eye Says It All: Periocular Region Methodologies. IEEE (2012)
15. Sadat, R.F.N., Rakib, A., Salehin, M.M., Afrin, N.: Efficient design of local binary pattern for image retrieval. IEEE Symposium on Computers and Informatics (2011)
16. Woodard, D.L., Pundlik, S.J., Lyle, J.R., Miller, P.E.: Periocular Region Appearance Cues for Biometric Identification. IEEE (2011)
17. Bharadwaj, S., Bhatt, H.S., Vatsa, M., Singh, R.: Periocular Biometrics: When Iris Recognition Fails. IEEE (2010)
18. UBIPr home page. https://iris.di.ubi.pt/ubipr.html
19. VSIOB home page. https://umkc.app.box.com/s/uwsynywq1ox9bb613bej3taw8t62okbe

Chapter 32
RPCR: Retransmission Power-Cost Aware Routing in Mobile Ad Hoc Networks



V. Sowmya Devi and Nagaratna P. Hegde

Abstract This paper presents a power control routing in MANETs to optimize the utilization of battery power of mobile nodes. The proposed approach considers a power required to transmit information on a particular path termed as power status and the amount of energy required for a node to send the information to its neighboring nodes, termed as node cost status. It also considers the probability of link error for each and every path, thus termed as retransmission power-cost aware routing (RPCR). The link error probability gives the information about the power required for retransmission. Thus, by including link error probability during energy evaluation of a particular path, the average energy consumption will be reduced. The performance of the proposed RPCR was verified.

32.1 Introduction

Mobile Adhoc wireless network is a group of mobile devices with networking capability and interfaces. It is self-organizing and adaptive in nature [1]. Power management [2] is the biggest challenge in mobile ad hoc networks. The amount of power required will be increased with the increase in the data rate and also with distance. In recent days, the transmission of multimedia information through mobile nodes is getting increased. For a mobile node having less energy can't support to transmit the multimedia data. The sudden drop in the power of a node reduces the quality of service of the entire network. If the node consumes maximum power to transmit information, it will drain up quickly. Thus, an energy-efficient routing approach needs to be designed with the view of reducing transmission energy of individual nodes and simultaneously increasing the network throughput. By reducing the individual power of nodes for each transmission, the total lifetime of node is getting increased.

V. Sowmya Devi
Vignana Bharathi Institute of Technology, Hyderabad, India

N. P. Hegde (✉)
Vasavi College of Engineering, Hyderabad, India
e-mail: nagaratnaph@staff.vce.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_32

This paper presents a retransmission power and cost-aware routing (RPCR) approach in MANETs. In this approach, the source node selects the best possible path from a given source and destination by considering the power status of path and node energy. This approach also includes probability of link error during energy evaluation for a particular established route. By considering the link error probability, the average amount of energy required for a successful transmission can be reduced. Results of simulation prove that the suggested approach is efficient routing approach for MANETs. The paper is organized as follows: Sect. 32.2 gives the information about the literature review. Section 32.3 illustrates the complete particulars about the power and cost-aware routing. Section 32.4 deals with the new link cost metric. The new link cost metric developed in this paper considers the link error probability during path selection. Section 32.5 deals with the simulation results and finally Sect. 32.6 provides the conclusions.

32.2 Literature Review

The routing protocols in MANETs are generally categorized as proactive (table-driven) and reactive (on-demand), based on the updating time of routes. Many routing protocols including Destination-Sequenced Distance Vector (DSDV) [3] and Fisheye State Routing protocol [4] belong to proactive routing and they differ in the number of times the routing tables are manipulated and the approaches used to exchange and maintain routing tables. In contrast to proactive protocols, not all up-to-date routes are maintained at each node. Dynamic Source Routing (DSR) [5] and Ad-Hoc On-Demand Distance Vector (AODV) [6] are examples of on-demand protocols. On the other hand, to establish efficient routes between the nodes, one important goal of a routing protocol is to keep the network alive as long as possible. As discussed in the introduction, this can be achieved by reducing the energy of mobile nodes during active and inactive communication. Load distribution [7], Power-down/Sleep mode [8] is also used to reduce the power utilization of mobile nodes, but this technique is for inactive communication. In this approach, the node which is not contributing in the communication is kept in power down/sleep mode to reduce the total energy consumption. This approach was focusing on reduction of nodes energy indirectly but not on the nodes participating in the communication. Transmission power control techniques [9] provide control on the transmission power of individual mobiles nodes. The primary interest of these techniques is to decrease the amount of the power required for transmission at node level. The main goal of transmission power control routing approaches is not to make available energy-efficient routes but to make the given route with energy efficiency by regulating the transmission power that is sufficient to contact the next-hop node. In [10], a flow augmented routing approach was proposed to minimize the node transmission energy by considering initial and residual energies. However, this approach was built based on assumption of static network. One more issue, the routing approach assumes that Data generation rate is presented in advance during each time step at all nodes. To overcome this

problem, in [11] a min–max routing approach was proposed without knowing the data generation in beforehand. This focuses on two parameters, minimizing power consumption and maximizing the minimal residual power. In this approach, a perturbation method was used to evaluate the minimum power of all paths available for a particular source node. However, the routing was based on an assumption of constant traffic. In [12] a power-aware routing was proposed to optimize the power levels of nodes [12]. Completely focused on the power levels of nodes which are willing to transmit but not on the power levels of neighboring nodes.

32.3 Power-Cost Aware Routing

In this approach, the routing was done based on the power status and node cost status. The power status of node is defined as the amount of the power required to transmit information to its neighboring nodes. The node cost status is defined as the amount of the power required to transmit the information of neighboring nodes. The node which was willing to send information towards the destination will calculate the self-power status and also the node cost status. This was accomplished by knowing 1-hop neighbors of neighboring node. A direct communication may put away more energy when compared with indirect communication through in-between nodes with the super linear relationship among transmission distance and energy.

In Fig. 32.1, when node 'A' sends data packet to node 'D', it can either send directly to 'D' or through its neighbors (B, C or E). Here, A to its neighboring node (B or C or E) is a direct transmission while neighboring node (B or C or E) to D is an indirect transmission with few intermediate nodes between immediate neighboring node (B or C or E) and D. For choosing the best path, node A computes and compares the consumed power of each path candidate. Power consumption of the direct transmission, $p(d)$ can be computed if the distance is known, i.e. $p(d) \propto d$, where d is the distance among the two nodes. The total power consumption of each

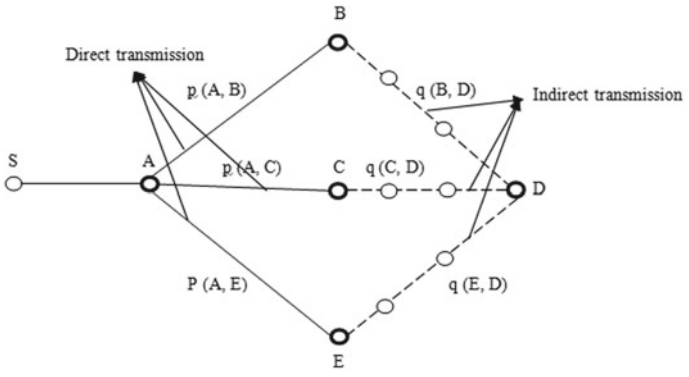


Fig. 32.1 Selection of the next hop node

and every path can be reduced by reducing the number of intermediate node. Let $q(d)$ be the amount of power required for data transmission between nodes (B or C or E) and D. Hence, node A, whether it is an intermediate or source node, chooses one of its neighbors (B, C or E) as the next hop which reduces $P(d) = p(d) + q(d)$.

In this approach, the source or any node which has information to send to destination will initiates route discovery and finds the possible paths. The route discovery phase was accomplished by flooding a RREQ packet with a destination ID. After receiving the RREQ, each and every node replies with RREP to source node. After this, the source node having N available paths to destination through various intermediate nodes. In every path, each and every node having information about previous hop node and next hop node. The source node which wants to send data has to select an optimal path among the available paths. For this purpose, in this approach, a power-cost aware routing was proposed. Initially, the source node knows his neighboring nodes and also the distance away from it. The super relationship $p(d) \propto d$ gives the information about the required power to transmit up to neighbor node. But it doesn't know about the 1-hop neighbors of its neighbor node. The source needs to find all nodes present in the available path. Based on the total power evaluated the source node finalizes the path. For this purpose, the source node sends a 1-hop REQ packet for each and every neighbor node of it, asking about the node cost status. The node cost status includes the available neighbor nodes, distance of neighbor nodes and power required for individual transmission to neighbor node. Then the neighbor node which received the 1-hop REQ replies with a 1-hop REP packet. Then the source node decides the optimal path which is having minimum power consumption.

In Fig. 32.1, let the power required for direct transmission be $p(d)$ and the power required for indirect transmission be $q(d)$, the path which is having minimum power consumption was selected as optimal path. Here, direct transmission refers to the transmission between source and its neighbor nodes, and indirect transmission refers to the transmission between neighbor node and its 1-hop neighbors. The amount of the power required will be varying. This is completely depending on the number of intermediate nodes present after first neighboring node. In the Fig. 32.1, the solid line represents direct transmission and dotted line represents indirect transmission. The holes on the dotted line represents intermediate node of the indirect transmission $q(d)$. Will vary with intermediate nodes on the indirect transmission. Due to the node mobility in MANETS, the intermediate nodes will not be constant. It also consumes more time, thus increases time delay. So, the proposed approach takes run-time decision according to the above discussion, i.e., the node won't perform all these function in addition during data transmission. It will update automatically in the background.

32.4 Retransmission-Power-Cost Aware Routing (RPCR)

The one more issue in MANETs, path breakages due to the mobility of nodes. In MANETs, the mobile nodes are moving in nature. So, the path established between

a source–destination pair may not be constant. The intermediate nodes on the established path may move out of range so the path breakage will occur at that node. In MANET, any breakage in link allows retransmission of information intern increases the energy consumption of the node and makes the node drain up quickly, because the amount of energy required for retransmission of data is equal to the energy consumed for sending the information before breakage [13]. Hence, there is a need to determine the link errors on the established path frequently. After the occurrence of link error, the node at which the error occurred needs to update to the source node. For this purpose, in [14], the intermediate node sends a packet denoting that the link error was occurred and also for re-route establishment. This scenario consumes more power at each and every node transmitting that packet.

In this approach, a new link cost-based routing approach was proposed to reduce the transmission energy by considering the effect of transmission link errors. This approach studies how link error rates affect this retransmission-aware metric. Consider a route from a source node S to a destination node D that has N-1 intermediate nodes named as 2, 3 … N (1 refers to the name of the source and $N + 1$ is the destination). The transmission energy over each link is $p_{i,i+1} = d_{i,i+1}$, where $d_{i,i+1}$ refers to the distance among nodes i and $i + 1$. Assuming that each of N links ($L_{1,2}$, $L_{2,3}$,.., $L_{\mathrm{N,D}}$) has an independent link error rate of $e_{i,i+1}$, the number of transmissions(including retransmissions) between node i and $i + 1$ is a geometrically distributed random variable as

$$Prob\{X = x\} = e_{i,i+1}^{x-1} \times \left(1 - e_{i,i+1}\right) \tag{32.1}$$

The mean number of transmissions for the successful transfer of a single packet is $1/\left(1 - e_{i,i+1}\right)$.Therefore, the effective transmission energy between nodes I and $i + 1$, which includes the effect of the transmission link error, is [14]

$$P_{i,i+1} = p_{i,i+1} \times 1/\left(1 - e_{i,i+1}\right) = ad_{i,i+1}^{\alpha} \times 1/\left(1 - e_{i,i+1}\right) \tag{32.2}$$

When the packet-error rate ($e_{i,i+1}$) is not negligible, the benefit of indirect transmission via intermediate nodes can be overshadowed by the inflation factor, $1/\left(1 - e_{i,i+1}\right)$. From the above equation, it is clear that, as the link error increases, the amount of power required for transmission is also getting increased. But compared to the power consumed for re-route establishment after occurrence of link error, the amount of power required for transmission by including a probability of link error is less. If there is any link error, the complete process has to take place again; the total power consumption will be simply doubled which will be high compared to power required with above power required using Eq. (32.2).

Algorithm: The complete proposed approach was done under two phases, route discovery and route maintenance.

Phase 1: Route Discovery.

Step 1: source node which has information to send to destination broadcasts a RREQ to all nodes in the network with destination ID.

Step 2: after receiving RREQ, each and every node replies with a RREP.

Phase 2: Route Maintenance.

Step 3: once the route request process over and the route was established, the destination node broadcasts route reply packet. If source node founds more number of paths, then it will choose a path having minimum energy consumption.

Step 4: The source node evaluates the amount of power required to transmit the information to just neighbor node by considering link error probability and sends 1-hop REQ packet to its just neighbor nodes to know the information (distance and power required) about their 1-hop neighbors.

Step 5: the intermediate nodes those are received 1-hop request from source node forwards it to their 1-hop neighbors and asks for their neighboring nodes. This will take very much time, so an elapsed time was fixed. The neighboring node replies with the received information within the elapsed time.

Step 6: upon receiving the information from just neighboring node through 1-hop REP packet, the source node selects a path which is having minimum energy consumption and transmits the data to the respective immediate neighbor node.

Step 7: This process continues until the destination was reached.

The above-illustrated process was completely done in the background. i.e., the node through which the data was processing already kept the entire information about its 1-hop neighbors and upon the receiving data from its previous node, it just chooses the node from the list and forwards the data. Note that the 1-hop REP was sent to the neighboring nodes not flooded.

32.5 Simulation Results

The performance of proposed approach was verified by varying the network parameter node mobility.

For the varied node mobility and the respective performance metrics, Packet Delivery Ratio, Average End-to-End delay, Average Energy Consumption, and Routing Overhead were evaluated. The proposed RPCR was compared with earlier Power Aware Routing (PAR) [15] and also with the basic AODV routing protocol to show the enhancement. In this, the number of nodes are 50, the pause time was fixed to 60 s and the node mobility was varied from 0 to 50 m/s.

From Fig. 32.2, it is clear that as the node mobility was varying, the packet delivery ration is decreasing. The proposed approach considered the link error probability during energy evaluation for all available paths. This reduces the probability of selecting a path with high link error. Thus, the packet delivery ratio is high. In Fig. 32.3, it is clear that as the node speed is increasing, the end-to-end delay is also increasing. In Fig. 32.4, it is clear that the average energy consumption was increasing with an increment in the node speed. In Fig. 32.5, as the node speed was increasing, the control packets were also increasing intern increasing the routing overhead.

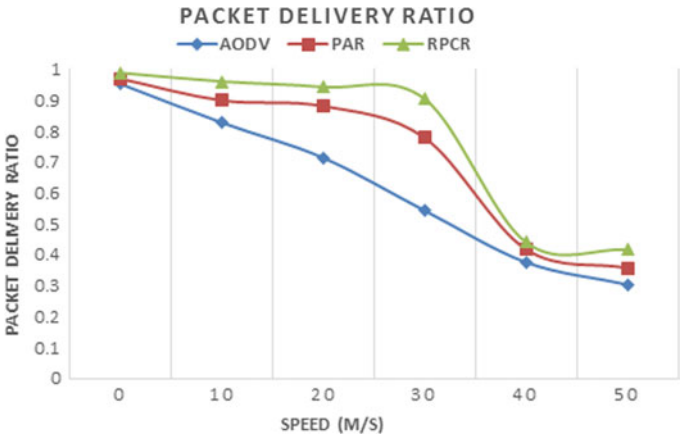


Fig. 32.2 Packet delivery ratio versus node mobility

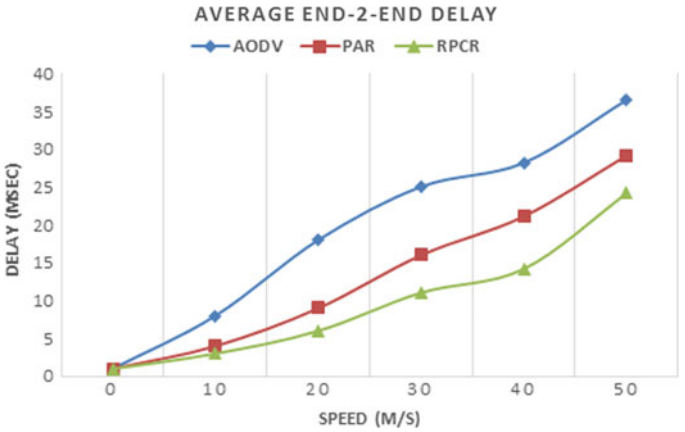


Fig. 32.3 Average end-to-end delay versus node mobility

32.6 Conclusion

This paper proposed a retransmission power and cost-aware routing in MANETs. In the proposed approach, the complete routing was done by considering node cost as well as path power cost, i.e., the amount of remaining energy at each and every node and the amount of power required for a particular available. The source node initially obtains few available paths by sending RREQ with destination ID. After getting available paths, the source node selects an optimal path by considering the path cost and node cost. Here, the path cost considers the link error probability. The simulation results are drawn by varying node mobility, the proposed approach was verified through packet delivery ratio, average end-to-end delay, average energy

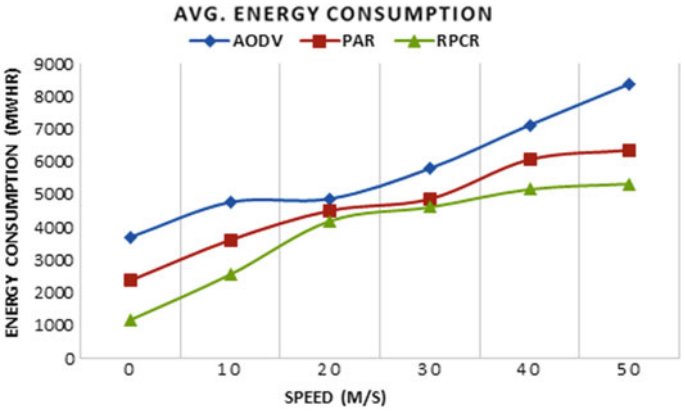


Fig. 32.4 Average energy consumption versus node mobility

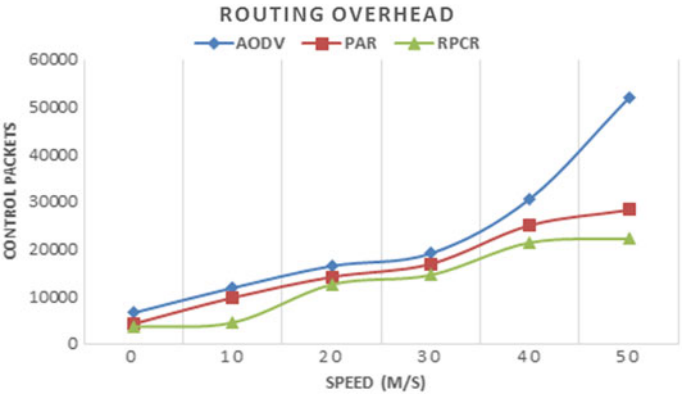


Fig. 32.5 Routing overhead versus node mobility

consumption, and routing overhead. Finally, the simulation results proved that the proposed approach is efficient in MANETs for power optimization.

References

1. Perkins, C.: Ad Hoc Networking, pp. 1–28. Addison-Wesley, Reading, MA (2001)
2. Narayana Swamy, S., Kawadia, V., Sreenivas, R.S., Kumar, P.R. Power control in ad-hoc networks: theory, architecture, algorithm and implementation of the COMPOW protocol. In: Proceedings of European Wireless 2002, pp. 156–162
3. Perkins, C., Bhagwat, P.: Highly dynamic destination sequenced distance-vector routing (DSDV) for mobile computers. Comput. Commun. Rev. **24**(4), 234–244 (1994)

4. Pei, G., Gerla, M., Chen, T.-W.: Fisheye state routing: a routing scheme for ad hoc wireless networks. In: Proceedings of IEEE International Conference on Communications (ICC), pp. 70–74 (2000)
5. Johnson, D., Maltz, D.: Dynamic source routing in ad hoc wireless networks. In: Imielinski, T., Korth, H. (eds.) Mobile Computing, pp. 153–181. Kluwer Academic (1996)
6. Perkins, C., Royer, E.: Ad-hoc on-demand distance vector routing. In: Proceedings of 2nd IEEE Workshop on Mobile Computing Systems and Applications, pp. 90–100 (1990)
7. Woo, K., Yu, C., Youn, H.Y., Lee, B.: Non-blocking, localized routing algorithm for balanced energy consumption in mobile ad hoc networks. In: Proceedings of International Symposium on Modeling Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2001), pp. 117–124 (2001)
8. Chen, B., Jamieson, K., Morris, R., Balakrishnan, H.: Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. In: Proceedings of International Conference on Mobile Computing and Networking (MobiCom'2001), pp. 85–96 (2001)
9. Stojmenovic, I., Lin, X.: Power-aware localized routing in wireless networks. IEEE Trans. Parallel Distrib. Syst. **12**(11), 1122–1133 (2001)
10. Chang, J.-H., Tassiulas, L.: Energy conserving routing in wireless ad-hoc networks. In: Proceedings of the Conference on Computer Communications (IEEE Infocom 2000), pp. 22–31 (2000)
11. Li, Q., Aslam, J., Rus, D.: Online power-aware routing in wireless ad-hoc networks. In: Proceedings of International Conference on Mobile Computing and Networking (MobiCom'2001), pp. 97–107 (2001)
12. Baisakh: Energy saving and survival routing protocol for mobile ad hoc networks. Int. J. Comput. Appl. (0975–888) **48**(2), (2012)
13. Sowmya Devi, V., Hegde, N.P.: Reliable and power efficient routing protocol for MANETs. Int. J. Comput. Netw. Inf. Secur. (IJCNIS) **8**(10), 61–69 (2016). https://doi.org/10.5815/ijcnis.2016.10.08
14. Hossain, M.J., Ali Akber Dewan, M., Chae, O.: Maximizing the effective lifetime of mobile ad hoc networks. IEICE Trans. Commun. **91**(9), 2818–2827 (2008)
15. Arvind, S., Zhen, L.: Maximum lifetime routing in wireless ad-hoc networks. In: INFOCOM 2004. Twenty-Third Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 2, pp. 1089–97. Hong Kong, China (March 2004)

Chapter 33
Comparison on Different Filters for Performance Improvement on Fingerprint Image Enhancement



Meghna B. Patel, Jagruti N. Patel, Satyen M. Parikh, and Ashok R. Patel

Abstract Today, in the era of Internet, biometric authentication plays many important roles. And for biometric authentication image, enhancement plays a vital role because the recognition of an image is extremely depend on it. The image quality can be improved using image enhancement phase after removing noise, joining damaged ridges and constructing smooth images. Various filtering techniques for image enhancement like min filter, max filter, median filter, mean filter, Gabor filter, O'Gorman filter as well as segmentation techniques like Sobel edge detection are used. In the proposed research work, above all mentioned filters are applied and compared the result for enhancing the fingerprint images. The comparison of all filters is done based on MSE, PSNR and computational time to measure the performance of the proposed research work. The implementation of the proposed research work was done using Java and used custom database of 100 images of 25 users.

33.1 Introduction

Image enhancement is the best process for result improvements for creating and increasing image appearances in the field of image processing. In real word, the value of the images is affected and becomes weak during the transmission from camera, level of light and sensor. The finger images become weak because of dry

M. B. Patel (✉) · S. M. Parikh
A.M. Patel Institute of Computer Studies, Ganpat University, Gujarat, India
e-mail: meghna.patel@ganpatuniversity.ac.in

S. M. Parikh
e-mail: parikhsatyen@yahoo.com

J. N. Patel (✉)
Department of Computer Science, Ganpat University, Gujarat, India
e-mail: jnp01@ganpatuniversity.ac.in

A. R. Patel
Florida Polytechnic University, Lakeland, FL, USA
e-mail: apatel@floridapoly.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_33

skin, sensor noise, cuts and scare, different pressure on sensor which generate non-adaptive noise as well as add nonlinearities among ridges point and valley design of impression [1]. The image quality improvement can be done using the enhancement technique that work for two bases: to join and damage ridges as well as eliminate noise among the ridges plus increase the ridges point valley disparity. To remove this type of noise, various filters like min–max filter, mean filter and median filter are used in image processing. These all filters are used as window masks for filtering. The filter is a minor grouping defined by kernel applied to every pixel and its neighbour's pixel of image. The filtering operation is worked based on region process, in that the importance of particular output pixel is resolute by applying an algorithm on value of the pixels in the region of the equivalent input pixel. A pixel area is a set of pixels defined by their position relative to that pixel. In most applications, the centre of the kernel is aligned with the current pixel and is a square with an odd number of elements (3, 5, 7, etc.) in each dimension. The technique used to apply a filter to an image is called convolution and can be applied in the spatial or frequency domain. In the spatial domain, if the kernel is pixel-centred, most of the convolution procedures multiply the kernel element with the corresponding pixel value. The elements of the resulting group are averages, and then the unique pixel standards are changed using the outcome.

33.2 Related Work

Image enhancement of fingerprint images can be able to perform either on gray scale or binary images [2, 3]. Binarization of the process will take place prior to image improvement, which misplaces the value of unique fingerprint data and create extra spurious minutiae structure. This drawback produces a problem in the objects improvement stage. An alternative method to use for object improvement is on a grey scale image which estimates the reliable local frequency of ridges and estimates the orientation.

There are three main methods to improve object quality: histogram flattening, FFT transform and Gabor filter [4]. The first process uses the normalization method at the primary stage using mean and variance. This is a pixel-based improvement technique that expands the legibility of the thumbprint and then does not change the ridge construction [5, 6]. The additional transformation process is established on FFT; improves the object quality that depends on the frequency domain. The last process, the Gabor filter, enhances the object quality according to its frequency and orientation. Throughout a collected work analysis [7–10], it was found that the maximum technique used in image enhancement is built using context filters. It depends on local orientation and peak frequency parameters. Filters were defined in Fourier or spatial domain. Conventional filtering works by using a single filter are to convolve the entire image. The filter characteristics vary depending on the local context [11, 12].

Gabor Filter is applied to research papers [7, 13] and proves that compared to histogram equalization and FFT-based enhancement, Gabor filter is efficient and improves the result of fingerprint images. Gabor filter has optimum joint solution for both frequency and spatial domains as well as has orientation and frequency selective properties. Secondly, the applied contextual filter is for image enhancement given in [14]. The researcher used a kernel, the major axis of which is oriented parallel to the ridges, called anisotropic smoothing.

A research paper [8] proposed an integration model for improving fingerprint images. In their work, the break at the two ends of the broken edges is effectively joined. As a result, two bogus ending trivia have been removed. Noise is suppressed in small and medium valleys. Edge holes have been completely removed, and borders have been improved as well. However, the proposed model does not solve some problems. When there is a high level of noise in the image, it is difficult to evaluate the positioning field in the object. In this regard, the segmentation of the original image is necessary. The performance of the proposed work is achieved by using the difference in the direction of the blocks and the quality index of small parts extracted by grouping the image quality characteristics. The proposed method allows to improve the difference in the directions of the blocks and the quality index and also allows to reduce the required time within reasonable limits. But the characteristic image factors for the identification and verification system still need to be improved. The search result still does not match the fingerprint.

The researcher [15] is working on new approaches for the confirmation of details in fingerprint images. They suggest that the new features for the details of fingerprint image confirmation provide better accurately than other grey scale methods stated in the literature. The methods are computationally effective and can also be used to design minutia detectors that can work straight on grey scale images. But merging the decision of the two classifiers and studying the effects of checking the details about the performance of the match are still needed to improve the object, and can be possible with the help of image improvement. Techniques based on direct grey scale enhancement work better than approaches that require binarization and thinning as intermediate steps [16]. The percentage of errors in terms of lost, exchanged and spurious little things that occur when using the binarization method is significantly lower than errors that occur when using these approaches. The modified Gabor filter works best, especially for poor images with low quality ridges and blocks with singular points. The requirement to assess the local frequency information performed by the Gabor filter is removed by using a single anisotropic filter. Improvement always requires speed and effectiveness.

Researcher [17, 18] used an incline approach to compute that direction path was orthogonal to the incline. First, we divided the object into square blocks and calculated the incline for each pixel in the x and y directions. The direction vector for each block is calculated. This is done by execution of an averaging process on all paths in the block that are orthogonal to the incline pixels. Use a low-pass filter to smooth the image because the ridge orientation may not be detected completely due to the noise in the image.

33.3 Proposed Research Work

Various filtering methods are available in many research papers for removing noise. Mainly, it is divided into two types: (1) linear filtering techniques and (2) nonlinear filtering techniques. In linear filtering methods, the output remains linear as per its inputs, but in nonlinear, the output is not remaining linear.

33.3.1 Mean Filter

This is a kind of linear spatial filter. It alters the values of noisy pixels by a regular value of its neighbourhood pixels containing this one [19, 20]. It is normally used for smoothing the images [1]. It removes unrepresentative pixels of their surroundings. To calculate mean, this filter used kernel to represent size and shape of neighbourhood. Most of 3×3 size kernels are used. Like Eq. (33.1), the kernel A is used with size $m = 3$, then

$$A_{\text{avg}} = \frac{1}{3}[1 \ 1 \ 1] \tag{33.1}$$

33.3.2 Min Filter

It is one type of linear filter. It is also known as zero percentile filter. It alters the value of noisy pixel by minimum intensity value among all neighbourhood pixels. It is used for finding out blackest point in image and improves the black area using the minimum value of the image. The min filter is defined by Eq. (33.2).

$$f^{\wedge(x,y)} = \min\{g(s,t)\} \ \text{ where } (s,t) \in s(x,y) \tag{33.2}$$

33.3.3 Max Filter

It is one type of linear filter. It is called as 100% filter. It alters the value of noisy pixel by maximum intensity value among all neighbourhood pixels [19]. This filter finds out optimistic points in image. It improves the bright areas of the image by finding out maximum value in the surrounded area by filter. It is mostly used to remove dotted noise occurred due to lower intensity value occurred because of maximum operations [21]. The max filter is defined by Eq. (33.3).

$$f^{\wedge(x,y)} = \max\{g(s,t)\} \text{ where } (s,t) \in s(x,y) \tag{33.3}$$

33.3.4 Median Filter

It is one type of nonlinear filter and also called as order statistical filter. It is used as a mean filter to remove noise by smoothing the image. It also decreases the intensity values among pixels. It alters the noisy pixel by the median value of its neighbourhood pixels. It is better compared to mean filter to preserve useful information in image [21]. The median filter is defined by Eq. (33.4).

$$(m,n) = \text{median}\{[i\ \ j],\ [i\ \ j] \in w\} \tag{33.4}$$

where w indicates the neighbourhood to the nearby location.

33.3.5 Sobel Edge Detection [10, 22]

The measurement of the two-dimensional spatial gradient in the object is performed by the Sobel operator. It is used to get the complete value of the approximate gradient at every pixel of the input grey scale image. Use two 3x3 convolution masks, one mask for to evaluate the gradient in the x route (columns) and the second one used to evaluate the gradient in the y(row) direction. The 3×3 matrix is much smaller than the real object. For this reason, the matrix moves across the object and simultaneously uses the four sides of pixels. The illustration of real Sobel covers is assumed under Fig. 33.1.

The estimated value is considered with the resulting equivalence (33.5):

$$|G| = |Gx| + |Gy| \tag{33.5}$$

The matrix is applied to the part of the input object, changes the pixel worth, changes single pixel in the right path, and then rests in the right path until it reaches the last of the line. Then, it drives to the next row.

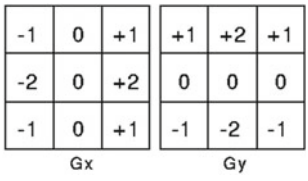


Fig. 33.1 Sobel masks illustration [23]

33.3.6 Gabor Filter [13, 24, 25]

In thumbprint identification for thumbprint identical and thumbprint arrangement, the Gabor filter is commonly defined as image improvement. The finger image covers dual key characteristics, such as the frequency and orientation of local peaks. The Gabor filter contains selective frequency and orientation properties, which is why it works most effectively for improving fingerprint. The Gabor filter is useful to every pixel situation in the whole image based on these properties. The outcome is used to filter and improves edges concerned with the path of native location and reductions whatever that is concerned with another way. Therefore, the filter improves ridge structure and smoothen valleys values, and simultaneously also effectively reduces noise. The following Eqs. (33.6) and (33.7) show the common method of Gabor filter [26].

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right) \tag{33.6}$$

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \tag{33.7}$$

$$x' = x \cos\theta + y \sin\theta$$

$$y' = -x \sin\theta + y \cos\theta$$

Here, the wavelength of sinusoidal factor is denoted as λ, angle of standard to similar lines of a Gabor purpose denoted as θ, stage offset denoted as ψ, sigma of the Gaussian packet denoted as σ and spatial aspect ratio denoted as γ.

33.3.7 O'Gorman Filter [13]

As per literature survey, the researcher states that O'Gorman filter also provides the good outcome for image enhancement and suits for fingerprint images [13, 14, 27, 28]. This is a directional filter. Related to the Gabor filter, the O'Gorman filter must use the pixel direction to create the filter, but it does not need the occurrence. To create this type of direction filter, first recognize a 7 * 7 matrix in a 0° direction and formerly revolve the filter base to the orientation of each pixel. Then, just calculate the rotation filter coordinates and determine the weight of the revolved filter of the 0° direction filter based on the rotation coordinate. The rotation methods have the classic rotation method as per Eq. (33.8):

$$\begin{pmatrix} i' \\ j' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} \tag{33.8}$$

A problem that comes with the formula of rotation is that the i' and j' do not have an integer value. For the sake of it, we have to make sure that the two pointers prove to be 0° in the direction filter and can be solved by interpolation and they can make other messengers through lined exclamation and establish the equivalent weightiness that corresponds to 0° filter.

33.4 Evaluation Parameter and Result Discussion

The implementation of proposed research work is done in JAVA platform and to check the performance of proposed image enhancement filters are measured using a custom database of 100 images of 25 users.

33.4.1 Quality of an Image

This evaluation parameter is used to check the measurement of visual appearance of an image. It checks the improvement in image quality after applying all the filters. Figure 33.2 shows the result of different filters in visual appearance. It proves that enhanced O'Gorman filter improves the result of an image after removing the noise and connects broken valleys and ridges compared to other filters.

33.4.2 PSNR and MSE Value

These two parameters are used to show the comparison of implemented different image enhancement filters. PSNR is well defined as the topmost signal-to-noise proportion among dual images using Eq. (33.9). Quality is measured from the inventive image and reinvented using it. The quality of the reinvented image is determined by the high PSNR and low MSE values.

MSE is well defined as the mean square error that approximates the alteration among the inventive image and the faded image. Example: If two images are indistinguishable on all sides, the MSE between them is measured as zero. First, compute the MSE value using an Eq. (33.9) and then use the following formula to calculate the PSNR value (33.10).

$$\text{MSE}\,(x, y) = \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2 \tag{33.9}$$

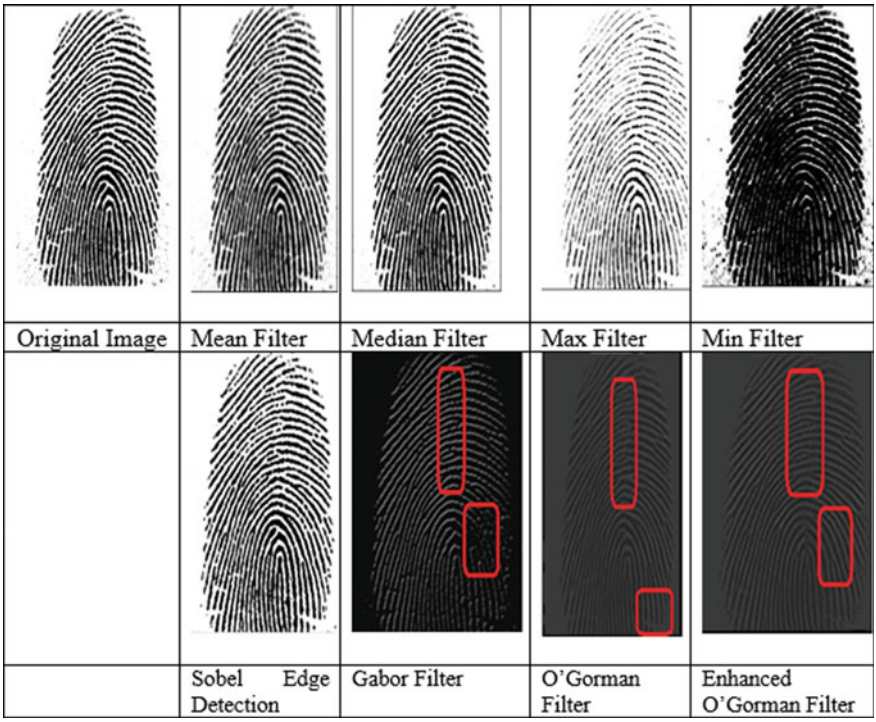


Fig. 33.2 Result of visual appearance of an image

Table 33.1 Result of PSNR value

Images of FingerDos databases	Mean filter	Median filter	Max filter	Min filter	Enhanced Sobel edge filter	Gabor filter	O'Gorman filter	Enhanced O'Gorman filter
0101.bmp	29.83	32.76	25.56	24.43	36.27	36.84	37.62	40.39
1201.bmp	29.93	32.96	24.34	23.82	36.92	37.06	37.96	40.81
3702.bmp	29.81	32.34	25.23	23.34	36.52	36.70	37.41	40.12
0401.bmp	29.47	32.45	24.94	234.13	36.72	36.94	37.76	40.19
1201.bmp	29.85	32.56	25.45	23.87	36.87	37.07	37.97	40.29
3901.bmp	29.67	32.43	25.78	24.36	36.24	36.80	37.56	40.30

$$\text{PSNR} = 10 \log_{10} \frac{L^2}{\text{MSE}} \tag{33.10}$$

where N is the total number of pixels in the input image and L is the number of discrete grey level.

Table 33.2 Result of MSE value

Images of FingerDos databases	Mean filter	Median filter	Max filter	Min filter	Enhanced Sobel edge filter	Gabor filter	O'Gorman filter	Enhanced O'Gorman filter
0101.bmp	16.23	15.54	18.43	19.12	13.13	13.45	12.61	5.94
1201.bmp	16.46	15.66	18.69	19.34	12.65	12.77	11.97	5.39
3702.bmp	16.52	14.92	17.07	18.87	13.23	13.89	13.03	6.32
0401.bmp	16.78	15.61	17.53	19.65	12.66	13.15	12.33	6.21
1201.bmp	17.59	14.28	16.32	19.48	12.72	12.74	11.94	6.08
3901.bmp	16.83	15.34	16.12	19.71	13.43	13.56	12.71	6.05

Tables 33.1 and 33.2 show the comparison of different filters using PSNR and MSE values. The comparison result of PSNR value of mean, median, max and min shows that the value of median filter is higher as well as MSE value of median filter is lower compared to other three filters. The PSNR and MSE value of Gabor filter and enhanced Sobel edge filter is nearer similar but better than median filter. And compared to Gabor filter and enhanced Sobel edge filter, the result of O'Gorman gives better result in terms of PSNR and MSE value and makes smooth images, but the connection of ridges is not done properly. So, enhance the O'Gorman filter with the use of interpolation which is used to connect broken ridges. The result of comparison proves that PSNR value of enhanced O'Gorman filter is higher and MSE value is lower compared to other filters which prove that the proposed algorithm gives a better result.

33.4.3 Computational Time

This evaluation parameter is used to compute the processing time of the enhancement process in milliseconds. Table 33.3 shows the summary of comparison between

Table 33.3 Result of computational time

Images of FingerDos databases	Mean filter	Median filter	Max filter	Min filter	Enhanced Sobel edge filter	Gabor filter	O'Gorman filter	Enhanced O'Gorman filter
0101.bmp	6	49	30	23	341	134	114	117
1201.bmp	7	48	31	22	344	137	114	114
3702.bmp	6	48	31	23	341	149	127	132
0401.bmp	6	49	30	23	339	143	122	124
1201.bmp	6	49	30	22	345	134	111	119
3901.bmp	7	48	30	23	349	143	124	122

different filters. In comparison with mean, median, max and min filter, the mean filter takes less processing time. And when we compare enhanced Sobel edge filter, Gabor filter, O'Gorman filter and enhanced filter, it proved proposed enhanced algorithm will take less time compared to other mentioned filters.

33.5 Conclusion

The proposed research work shows the implementation of eight different filters like mean, median, max, min, enhanced Sobel edge, Gabor, O'Gorman and enhanced O'Gorman for image enhancement. The evaluation of implemented proposed work can be compared using different standard parameters like PSNR, MSE, computational time and image quality measurement standard. The experimental study shows that the performance of the enhanced O'Gorman filter is better than other filters.

References

1. Tania, S., Rowaida, R.: A comparative study of various image filtering techniques for removing various noisy pixels in aerial image. Int. J. Signal Process. Image Process. Pattern Recogn. **9**(3), 113–124 (2016)
2. Maio, D., Maltoni, D.: Direct gray-scale minutiae detection in fingerprints. IEEE Trans. Pattern Anal. Mach. Intell. **19**(1), 27–40 (1997)
3. Patel, M.B., Parikh, S.M., Patel, A.R.: Performance improvement in binarization for fingerprint recognition. IOSR J. Comput. Eng **19**(3), 68–74 (2017)
4. Omran, S.S., Salih, M.A.: Comparative study of fingerprint image enhancement methods. J. Univ. Babylon **22**(4), 708–723 (2014)
5. Greenberg, S., Aladjem, M., Kogan, D.: Fingerprint image enhancement using filtering techniques. Real-Time Imaging **8**(3), 227–236 (2002)
6. Patel, M.B., Parikh, S.M., Patel, A.R.: Global normalization for fingerprint image enhancement. In: International Conference on Computational Vision and Bio Inspired Computing, Sept 2019, pp. 1059–1066. Springer, Cham
7. Hong, L., Wan, Y., Jain, A.: Fingerprint image enhancement: algorithm and performance evaluation. IEEE Trans. Pattern Anal. Mach. Intell. **20**(8), 777–789 (1998)
8. Wu, C., Shi, Z., Govindaraju, V.: Fingerprint image enhancement method using directional median filter. In: Biometric Technology for Human Identification, Aug 2004, vol. 5404, pp. 66–75. International Society for Optics and Photonics
9. Yang, G.Z., Burger, P., Firmin, D.N., Underwood, S.R.: Structure adaptive anisotropic image filtering. Image Vis. Comput. **14**(2), 135–145 (1996)
10. Patel, R.B., Patel, M.B., et al.: Performance improvement in fingerprint image enhancement using Gaussian mask and Sobel convolution. In: 10th International Conference on Transformation of Business, Economy and Society in Digital Era (2019)
11. Patel, M.B., Parikh, S.M., Patel, A.R.: Performance improvement in gradient based algorithm for the estimation of fingerprint orientation fields. Int. J. Comput. Appl. **167**(2), 12–18 (2017)
12. Patel, M., Parikh, S.M., Patel, A.R.: An improved approach in core point detection algorithm for fingerprint recognition. In: Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT), Apr 2018, pp. 26–27

13. Patel, M.B., Patel, R.B., Parikh, S.M., Patel, A.R.: An improved O'Gorman filter for fingerprint image enhancement. In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Aug 2017, pp. 200–209. IEEE
14. Patel, M.B., Parikh, S.M., Patel, A.R.: An improved approach in fingerprint recognition algorithm. In: Smart Computational Strategies: Theoretical and Practical Aspects, pp. 135–151. Springer, Singapore (2019)
15. Chikkerur, S., Govindaraju, V., Pankanti, S., Bolle, R., Ratha, N.: Novel approaches for minutiae verification in fingerprint images. In: 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05), Jan 2005, vol. 1, pp. 111–116. IEEE
16. Patel, M.B., Parikh, S.M., Patel, A.R.: An improved thinning algorithm for fingerprint recognition. Int. J. Adv. Res. Comput. Sci. **8**(7), 1238–1244 (2017)
17. Halici, U., Jain, L.C., Erol, A.: Introduction to fingerprint recognition. In: Intelligent Biometric Techniques in Fingerprint and Face Recognition, pp. 1–34 (1999)
18. Nie, P., Geng, W.B.: An introduction to fingerprint recognition technology. Comput. Knowl. Technol. (Acad. Exch.) **17** (2007)
19. Kaur, R., Singh, E.R.: Image filtering techniques—a review. Int. J. Adv. Res. Sci. Eng. **6**(08), 2066–2071 (2017)
20. Patel, N., Shah, A., Mistry, M., Dangarwala, K.: A study of digital image filtering techniques in spatial image processing. In: Proceedings of the 2014 International Conference on Convergence of Technology (I2CT), pp. 1–6 (2014)
21. Rani, K.S., Rao, D.N.: A comparative study of various noise removal techniques using filters. J. Eng. Technol. **7**(2), 47–52 (2018)
22. Vincent, O.R., Folorunso, O.: A descriptive algorithm for Sobel image edge detection. In: Proceedings of Informing Science & IT Education Conference (InSITE), June 2009, vol. 40, pp. 97–107. Informing Science Institute, California
23. Mohammad, E., JawadKadhim, M., Hamad, W.I., Helyel, S.Y., Alrsaak, A.A.A., Al-Kazraji, F.K.S., Hadeeabud, A.M.: Study Sobel edge detection effect on the ImageEdges using MATLAB. Int. J. Innov. Res. Sci. Eng. Technol. **3**(3), 10408–10415 (2014)
24. Dunn, D., Higgins, W.E.: Optimal Gabor filters for texture segmentation. IEEE Trans. Image Process. **4**(7), 947–964 (1995)
25. Patel, M.B., Parikh, S.M., Patel, A.R.: Performance improvement in preprocessing phase of fingerprint recognition. In: Information and Communication Technology for Intelligent Systems, pp. 521–530. Springer, Singapore (2019)
26. Kawagoe, M., Tojo, A.: Fingerprint pattern classification. Pattern Recogn. **17**(3), 295–303 (1984)
27. O'Gorman, L., Nickerson, J.V.: An approach to fingerprint filter design. Pattern Recogn. **22**(1), 29–38 (1989)
28. Patel, M.B., Parikh, S.M., Patel, A.R.: An approach for scaling up performance of fingerprint recognition. Int. J. Comput. Sci. Eng. **7**(5), 457–461 (2019)

Chapter 34
Framework for Spam Detection Using Multi-objective Optimization Algorithm



M. Deepika and Nagaratna P. Hegde

Abstract Electronic spam is a big problem for huge international corporations like AOL, Google, Yahoo and Microsoft. Spam creates many issues and may cause economic damage. Spam causes many problems. Spam creates issues in traffic and bottlenecks, reducing memory, processing capacity and distance. We can describe classification output optimization in computer vision as a multi-objective optimization problem. This paper presents an evolutionary multi-objective optimization algorithm (E-MOA) of the anti-spam filtering issue that discusses both e-mail classification requirements (FP and FN error rates) and e-mail classification times (minimisation). Test findings using a freely accessible model corpus have enabled us to draw significant aspects concerning both the effectiveness of rule-based classification filtering and the appropriateness of a spam filtering two-way classification system. This dataset also uses the DNN classification method, and this classification involves cross-validation. Finally, the e-mail spam classifier is defined based on error rate, accuracy and recall.

34.1 Introduction

In the everyday existence of people all over the world, the usage of Internet mailing systems has become indispensable. In addition, integrating e-mail with the new constantly linked mobile smart phones allows an easy yet effective way to remain in contact with certain individuals and to share documents effectively at all times. In effect, all instant messaging (IM) programs and e-mail are widely used. The basic distinction between common IM (including WhatsApp or GTalk) applications and the Internet mailing systems is, therefore, that they can only use consent control

M. Deepika (✉)
Research Scholar, Osmania University, Hyderabad, India

Assistant Professor, Neil Gogte Institute of Technology, Hyderabad, Telangana, India

N. P. Hegde
Professor, Department of CSE, Vasavi College of Engineering, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_34

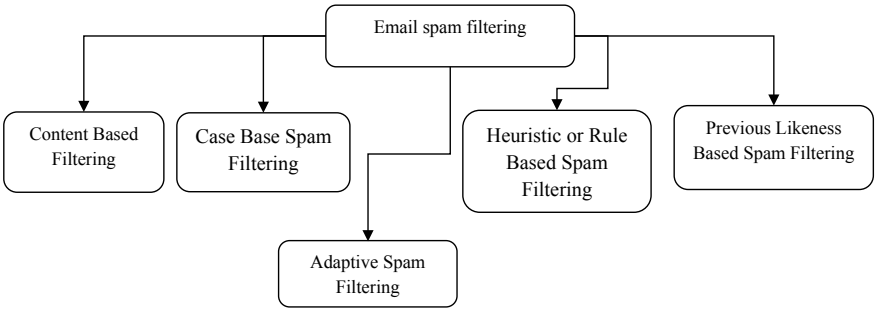


Fig. 34.1 E-mail spam filtering methods in existence

approaches within former applications. This made it possible to use e-mails as an assertive/huge means of advertisement and a virus delivery tool that causes spam. The latest figures [1, 2] suggest 40% of all communications spam, about 15.4 million communications a day, costing Internet users approximately 355 million dollars a year. In recent years, the filtering system [3] of spam assassin has been common employed in Web to reduce the spam. While there are many methods of e-mail spam filtering, state-of-the-art implementations are explored in this paper as seen in Fig. 34.1. Below are the various spam filtering strategies commonly used to solve the spam problem. This has established classification accuracy that could be optimized to test the advantages of using the evolutionary multi-objective optimization algorithm (E-MOA) to plan the multifunctional anti-spam filter scheduling issue.

Spam is no longer waste but is danger as it involves virus applications and spyware agents which destroy the recipient's network, which is why spam detection is developing. This has suggested much spam identification approaches based on machine learning algorithms. Through the usage of bulk mailing devices, the volume of spam has significantly risen, spam identification strategies will cope with it. For the identification of spam, it suggests security configurations and selection of features to minimize overhead operation, with high classification rates. However, subjective value and maximum number of features have not been taken into consideration in prior methods, so there are no strategies to using them together to date. In this article, we suggest an optimized random forest (RF) spam detection model to refine and select features. To improve detection speeds, we refine two RF parameters. We have the variable value of each function such that unnecessary features may be quickly omitted. In addition, we plan to use two approaches for optimizing the chosen features: (i) only optimizing one parameter in the overall feature range, and (ii) optimizing parameters in each removal step. We perform spambase dataset tests and explain the viability of our method.

34.2 Background Work

In [4], the authors presented stream clustering techniques for handling spam and in many circumstances categorizing inputs/tweets in spam and non-spam clusters. Such methods suggest that the cluster contains several neighboring (micro) clusters with an growing symmetrical micro-cluster distribution. However, this assumption is not necessarily valid, and small micro-clusters can be spread asymmetrically. To enhance the accuracy of previous online approaches, they recommend substituting the distance with a group of classifiers. These INBs should get the mean and limit of the micro-clusters, unreliable to the asymmetrical small micro-cluster. In this paper, DenStream promoted the alternative application INB-DenStream. The INB-DenStream data collection was productively applied by province-of-the-art approaches, including DenStream StreamKM++ and CluStream in terms of precise recall, F1-score and computational complexity. The quantitative outcomes revealed that their system is equal to adversaries in virtually all datasets.

In [5], the authors examined the popularity of the social media, but some attackers attracted their exponential growth. Many researchers have used spammers' behavior to reduce this possibility and have got promising outcomes by utilizing machine intelligence procedures to detect spammers on social networks. Many of them analyze disregard class imbalances of outcomes from the modern world. They introduce a complex stacking-based set learning framework in this article to improve the impact of class imbalances on the detection of spam on social networks. In the hybrid system, we incorporate cost-sensitive research into deep neural processing. They will combine the ingredients with the classification results for obtaining misclassification for prediction outcomes of the basic classifiers in complex terms. The experimental results show that the efficiency of spam detection for imbalanced datasets is growing inside our framework.

In [6], Twitter authors provide a range of special features; this cannot use directly traditional methods to Twitter spam research. The paper recommends a Twitter-only spam recognition program called TwitterSpamDetector. TwitterSpamDetector uses Twitter-specific features to recognize Twitter spam. The accuracy and response of TwitterSpamDetector are measured at 0.943 and 0.913, respectively, according to measurement observations. In [7], the authors spam e-mails are often a concern. It is, therefore, still very necessary to construct algorithms to solve this issue. This paper discusses two different algorithms for spam prevention. The first Bayesian-based algorithm, though, is improved if the Bayesian filter will not decide the data compression algorithm. The second approach uses the classification algorithm for the particle swarm optimization model. The results of the mentioned algorithms are promising.

34.3 Proposed Model

The overall design of the proposed system is given in Fig. 34.2, and each of these components is addressed in the following sections briefly.

Dataset: A dataset called the spambase UCI is accessible to test the proposed process. It comprises a variety of sections, including spam and non-spam. This dataset comprises 4601 instances and 57 attributes. The last column of the table shows whether the e-mail is a spam (1) (0). Many characteristics show that the e-mail is also used for a specific word or character. The runlength (55–57) attributes calculate the series length of successive capital letters.

Pre-processing: Most of the data in the physical world include composite, noisy and incomplete values. While it focuses the quality determination on the standard

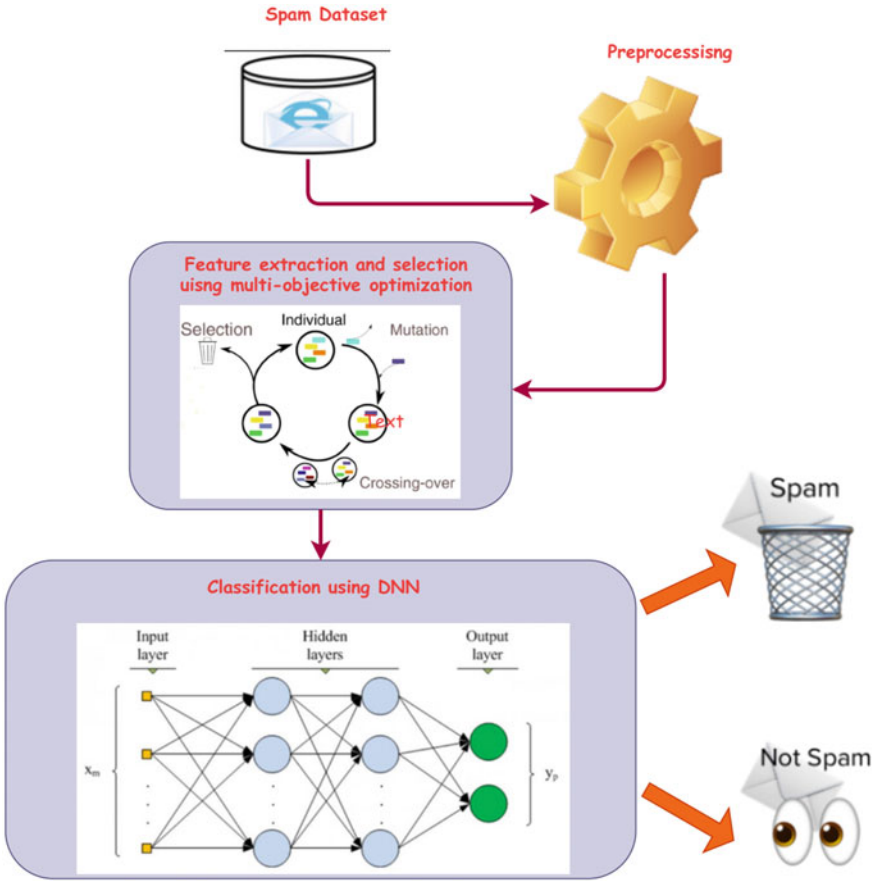


Fig. 34.2 Our proposed methodology for spambase UCI dataset classification

of mining, preprocessing is a very necessary activity well before it carries mining operation out. Data maintenance, data creation, data replication and data reduction are essential activities in the preprocessing of data. It achieves data normalization in this dataset until it performs the analysis.

Feature selection: Consider the dataset $\mathbb{Z}$ of dimension $m \times n$, where m is the number of features and n is the number of data points. Then, through point $z_{i,}$ marks the jth data point's ith function. Also let vector $\mathbb{L}$ denote the label of increasing data point of $\mathbb{Z}$, so that l_j denotes the jth data point. Here various attributes with the specific patterns should be noticed. Therefore, features within a set of [1, 10] are normalized as follows:

$$x_{i,j} = 1 + 9\left(\frac{Z_{i,j} - \min_{1\le k\le m} Z_{i,k}}{\max_{1\le k\le m} Z_{i,k} - \min_{1\le k\le m} Z_{i,k}}\right) \tag{34.1}$$

where $\mathbb{X}$ is a variable representing the intra-class distance from the practical point of view. This can find that there is a need to back up the intra-class distance from the intrinsic function to add the variable weight vector $\boldsymbol{W}$ at the end rather than identifying and w be the feature selection or weighting factor for the ith feature. This defines the difference between groups in the same way.

$$w_i = \begin{cases} 0, & \text{if the feature is rejected} \\ [1, \mathbb{A}], & \text{if the feature is selected} \end{cases} \tag{34.2}$$

$$Y_j = W \odot X_j \tag{34.3}$$

$$\begin{aligned} d(Y_p, Y_q) &= \sum_{i=1}^{m} |y_{i,p} - y_{i.q}| \\ &= \sum_{i=1}^{m} |w_i x_{i,p} - w_i x_{i.q}| = \sum_{i=1}^{m} w_i |x_{i,p} - x_{i.q}| \end{aligned}$$

Thus, $d(Y_p, Y_q) = W^T |X_p - X_q|$.
By Eq. (34.4), the total intra-class distance can be given as

$$\begin{aligned} D_{\text{intra}} &= \sum_{p=1}^{n} \sum_{q=p+1}^{n} {}_{\forall l_p = l_q} d(Y_p, Y_q) \\ &= W^T \sum_{p=1}^{n} \sum_{q=p+1}^{n} {}_{\forall l_p = l_q} |X_p - X_q| \end{aligned} \tag{34.4}$$

Thus,

$$D_{\text{intra}} = W^T \Delta^{\text{intra}} \tag{34.5}$$

where $\Delta^{\mathbf{intra}}$ is a variable representing the intra-class distance from the practical point of view. This can find it that there is a need to back up the intra-class distance from the intrinsic function to add the variable weight vector at the end rather than identifying and summing up the pair-way group from the data points. We may define the difference between groups in the same way.

$$\begin{aligned} D_{\text{inter}} &= \sum_{p=1}^{n} \sum_{q=p+1 \forall l_p \neq l_q}^{n} d\left(Y_p, Y_q\right) \\ &= W^T \sum_{p=1}^{n} \sum_{q=p+1 \forall l_p \neq l_q}^{n} \left|X_p - X_q\right| \\ D_{\text{inter}} &= W^T \Delta^{\text{inter}} \end{aligned} \tag{34.6}$$

To improve the affinity against the collection of applications, both intra-class/inter-class distances may apply to a penalty element. Under the definition of penalty features, the larger the amount of roles identified, the greater the degree to which it disciplines the target function. The target functions to be refined are then.

$$\begin{aligned} F_1(\boldsymbol{W}) &= D_{\text{intra}} + \lambda_1[\min\{\mathbf{1}, \boldsymbol{W}\}]^T \mathbf{1}, \\ F_2(\boldsymbol{W}) &= -D_{\text{inter}} + \lambda_2[\min\{\mathbf{1}, \boldsymbol{W}\}]^T \mathbf{1}, \end{aligned} \tag{34.7}$$

where λ_1 and λ_2 are penalty coefficients, with the help of F_1 and F_2 minimization to attain the optimal feature selection and weight vector.

$$\begin{aligned} &\text{Minimize } F(w) = [F_1(w), F_2(w)]^T \\ &\text{subjected to } W = [w_1, w_2, \ldots, w_m]^T \in 0 \cup [1, \mathbb{A}] \end{aligned} \tag{34.8}$$

Thus, Eq. (34.8) represents a multi-objective optimization problem (MOP).

Solving Multi-objective Optimization Problem: The optimum performance leads to a particular relation between candidate solutions, overcome by the Pareto-optimal principle where the best combination of such goal Pareto-optimal is obtained [8–11]. Then one solution is selected from the optimum front of Pareto, based on the problem.

In the proposed process, N function selection vectors (W) are initialized at the random level inside [0, A]. Then, MOEA/D requires a replication phase to slowly evolve weight vector population (W) toward optima. This process requires two phases in our system, namely mutation and crossover. It conducts both operations according to DE/rand/bin/1, an efficient single goal optimization algorithm, with variants of differential evolution (DE). After each reproduction process of MOEA/D, it performs a repair process over the freshly produced offspring to fulfill the optimization problem constraints. Remember that $\boldsymbol{W} = [w_1, w_2, \ldots, w_m]^T$ is the reproduction variable. Then the ith vector of $\boldsymbol{W}'$ dimension can be determined as follows.

$$w_i^{'} = \begin{cases} A, & \text{if } w_i > A \\ o, & \text{if } w_i < 1, \\ w_i, & \text{otherwise} \end{cases} \tag{34.9}$$

$\forall\, i \in [1, m]$. When the value for w_i is set to [0, 1), which is far less than its highest value A, the weight for the ith attribute is very small and can be called void. It achieves this in the second equation situation (34.9), raising the likelihood of the number of choices to decrease, thus by achieving a limit of μ^k, $\forall k$ provides best solution in case of Pareto. For obtaining the optimum N Pareto solutions given by MOEA/D resembles to a function ($\boldsymbol{W}_k$), the optimal vector ($\boldsymbol{W}^*$) is the following.

$$k^* = \max_k \mu^k,\ W^* = W^{k*} \tag{34.10}$$

To get the optimal feature subset, $\mathbb{W}^*$ and its related weights can be used to apply for classifier to detect the spam and non-spam.

Classification: In this paper, it is introduced with significant changes (DNN) which acts as classifier for class imbalance tasks. It is capable of learning their complex network architectures to retrieve higher-level functions [12–14]. They compose the DNN model of an input layer, hidden layer (s) and an output layer than seen in Fig. 34.3.

The input layer recognizes information from beyond the network. In the secret layer (s), multilevel input features are extracted linearly to partition various data forms. Hidden layer $h(h \in \{1, \ldots, H\})$ comprises a set of parameters $\theta_h = \{W_h, b_h\}$

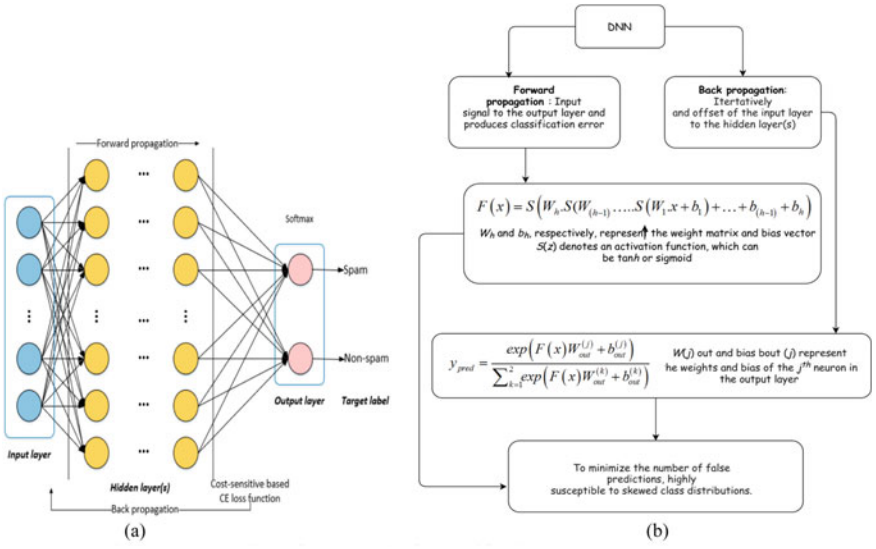


Fig. 34.3 **a** Overview of DNN model for classification, **b** flow process DNN model

in which W_h is a related weight matrix and b_h is a vector of prejudice. This method comprises two stages: forward propagation (FP) and back propagation (BP). The forward propagation sends the input signal to the output layer and induces an error in classification. This balances the hidden layer (s) of the input side, provided that DNN is completely connected with H hidden layers, as seen in Fig. 34.3, for an input function vector x, a complex feature transformation mechanism is represented in H hidden layers of the DNN by computation.

The DNN model for classification as shown in Fig. 34.3a can be explained with the help of flow process. Figure 34.3b consists of forward propagation (FP) and back propagation (BP).

34.4 Results and Discussion

In this section, DNN classification is used to characterize the test data using the chosen and weighted attribute vectors. Therefore, to show the functionality of the proposed algorithm, we checked and plotted our algorithm on a spambase dataset called MATLAB 2017a with number of features used in each case for obtaining accuracy listed in Table 34.1 (Fig. 34.4).

Table 34.2 and Fig. 34.5 show that the DNN classifier using E-MOA outperforms compared to other algorithms for feature selection because it can significantly

Table 34.1 Comparison of different feature selection (FS) algorithms

Dataset	Avg. classification accuracy in %	
	Fuzzy rule-based FS	Proposed
Spambase	93.24	98.12

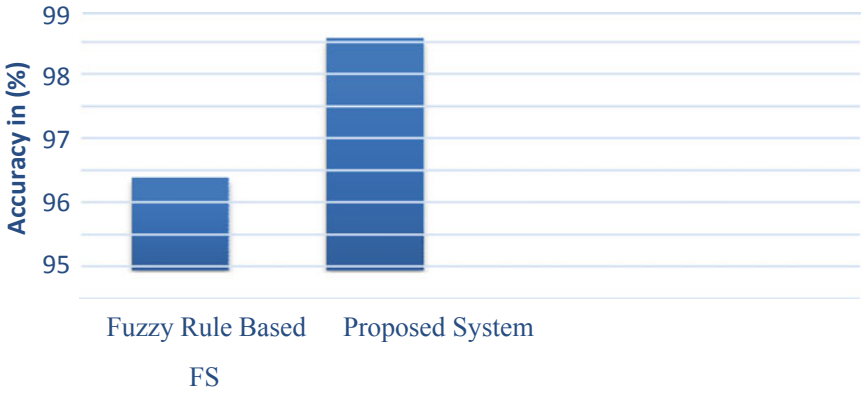


Fig. 34.4 Comparison of different feature selection algorithms with classification accuracy

Table 34.2 Results of classification approaches

Classification algorithm	Error rate before feature selection	Error rate before filtering		
		E-MOA	Genetic algorithm	Runs filtering
Naïve Bayes continuous	0.1136	0.1265	0.1328	0.1488
ID3	0.0896	0.0756	0.0789	0.0847
SVM	0.0375	0.0232	0.0256	0.0325
Our proposed DNN	0.0121	0.0079	0.0089	0.0093

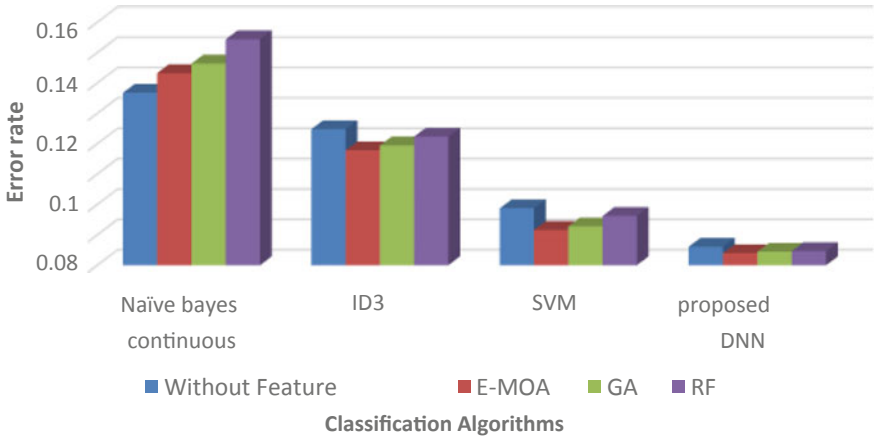


Fig. 34.5 Comparison of different feature selection algorithms with classification error rates

reduce dimensionality reduction of features in the E-MOA, contributing to less DNN classifier training information (Table 34.3).

The terms, positive and negative, are the estimation of the classifier, and the terms, true and false, apply to the prediction of the classifier. The DNN classification accuracy and recall are performed, and we show the results in Fig. 34.6. We regard the DNN classification as the best classifier based on the data, as it obtained 99% accuracy by selecting the E-MOA feature.

Table 34.3 Results of DNN classifier

Values prediction	Accuracy	Recall
Spam	0.9985	0.9905
Not spam	0.9842	0.9924

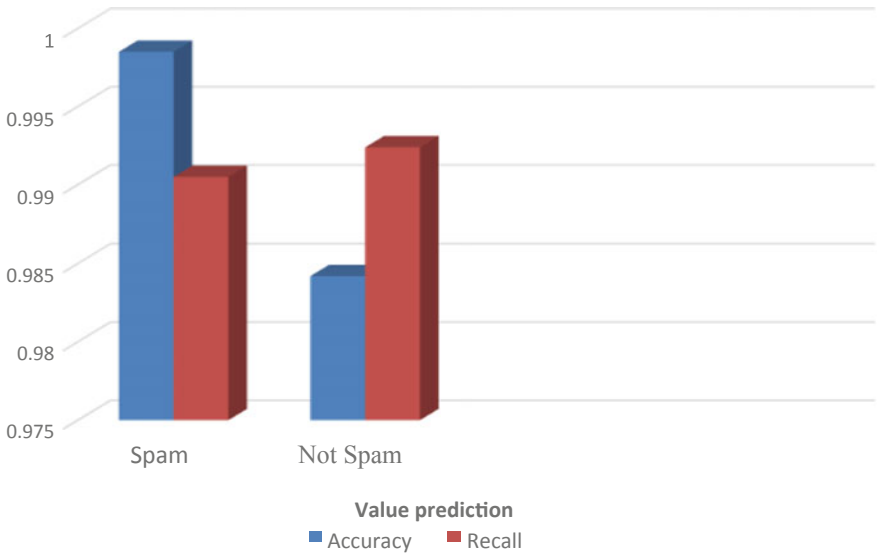


Fig. 34.6 Accuracy and recall of the spam classification algorithm (DNN)

34.5 Conclusion

In this study, the model has tested the usage of evolutionary, multi-target optimization algorithms to improve the consistency of the detection and classification of anti-spam filters (i.e., performance and efficacy of anti-spam filters). In addition, the E-MOA method offered strategies for ideal anti-spam filter settings that address many conditions for use of anti-spam filters. To remove issues such as bandwidth reduction and very poor performance that are suffering from, we may model the suggested algorithm to be applied on the mail server and mail application. The DNN classification algorithm applied on relevant features after E-MOA has resulted more than 99% accuracy, 99.01% recall and 0.0079 of error rate in spam detection.

References

1. Statista: The Statistics Portal: Global spam volume as percentage of total e-mail traffic from January 2014 to September 2016, by month (2016). https://www.statista.com/statistics/420391/spam-email-traffic-share/. Accessed 14 Feb 2017
2. Digital Marketing Ramblings: 73 incredible e-mail statistics (2016). https://expandedramblings.com/index.php/email-statistics/. Accessed 14 Feb 2017
3. The Apache SpamAssassin Group: The first enterprise open-source spam filter (2003). https://spamassassin.apache.org/. Accessed 14 Feb 2017
4. Tajalizadeh, H., Boostani, R., Tajalizadeh, H.: A novel stream clustering framework for spam detection in Twitter. IEEE Trans. Comput. Soc. Syst. 1–10 (2019). https://doi.org/10.1109/TCSS.2019.2910818

5. Zhao, C., Xin, Y., Li, X., Yang, Y., Chen, Y.: A heterogeneous ensemble learning framework for spam detection in social networks with imbalanced data. Appl. Sci. **10**, 936 (2020). https://doi.org/10.3390/app10030936
6. Kabakus, A.T., Kara, R.: "TwitterSpamDetector": a spam detection framework for Twitter. Int. J. Knowl. Syst. Sci. **10**, 1–14 (2019). https://doi.org/10.4018/IJKSS.2019070101
7. Lee, S., Kim, D., Kim, J., Park, J.: Spam detection using feature selection and parameters optimization. In: CISIS 2010—The 4th International Conference on Complex, Intelligent and Software Intensive Systems, pp. 883–888 (2010). https://doi.org/10.1109/CISIS.2010.116
8. Tian, Y., Peng, S., Rodemann, T., Zhang, X., Jin, Y.: Automated selection of evolutionary multi-objective optimization algorithms (2019). https://doi.org/10.1109/SSCI44817.2019.9003018
9. Gong, M., Jiao, L., Yang, D., Ma, W.: Research on evolutionary multi-objective optimization algorithms. J. Softw. **20** (2009). https://doi.org/10.3724/SP.J.1001.2009.03483
10. Saborido Infantes, R., Ruiz, A.B., Bermúdez, J., Vercher, E., Luque, M.: Evolutionary multi-objective optimization algorithms for fuzzy portfolio selection. Appl. Soft Comput. **39** (2015). https://doi.org/10.1016/j.asoc.2015.11.005
11. Cabrera, G., Vasconcellos, C., Soto, R., León, J.M., Paredes, F., Crawford, B.: An evolutionary multi-objective optimization algorithm for portfolio selection problem. Int. J. Phys. Sci. **6**, 5317–5328 (2011)
12. Ngamsuriyaroj, S., Taninpong, P.: Tree-based text stream clustering with application to spam mail classification. Int. J. Data Min. Model. Manag. **10**, 353 (2018). https://doi.org/10.1504/IJDMMM.2018.10015879
13. Sel, S., Hanbay, D.: E-mail classification using natural language processing, pp. 1–4 (2019). https://doi.org/10.1109/SIU.2019.8806593
14. Surwade, A.U., Patil, M., Kolhe, S.: An empirical analysis of spam E-mail classification using machine learning techniques. Int. J. Appl. Eng. Res. **9**, 5057–5074 (2014)

Chapter 35
Time Series Analysis of Assam Rainfall Using SARIMA and ARIMA



Utpal Barman, Asif Ekbal Hussain, Mridul Jyoti Dahal, Puja Barman, and Mehnaz Hazarika

Abstract Time series analysis of rainfall is very much essential for farming. Agriculture productivity is depended on rainfall. It is important to predict the future rainfall from farmers' point of view. In this paper, we apply seasonal auto-regressive integrated moving average (SARIMA) and auto-regressive integrated moving average (ARIMA) techniques for the monthly time series analysis of rainfall in Assam. The rainfall data contains the monthly rainfall of Assam from 1901 to 2017. Here, different components of the rainfall are visualized before apply the SARIMA and ARIMA. The handling procedures of seasonal components (p, d, and q) are reported using moving average, and augmented ducky fuller test. The ACF and PACF are used to find the seasonal components of the SARIMA and ARIMA. The SARIMA model is selected as the best model as compared to ARIMA based on AIC, BIC, HQIC, regression score (RC), mean absolute error (MAE), median absolute error (MeAE), mean squared error (MSE), mean squared log error (MSLE), and root mean square error (RMSE) of the analysis. The final results of the two methods are validated with the actual rainfall of Assam during the period.

35.1 Introduction

Rainfall is one major problem in Assam, and it indirectly affects the crop productivity in Assam. Every year heavy rainfall creates artificial flood in Assam. Due to that artificial flood, crop productivity in Assam is less as compared to other states of Assam. Agriculture productivity in India is primarily dependent on the rainfall. Nominal rainfall is always good for high agriculture productivity. But heavy rainfall is too dangerous for agriculture.

Researches are already done to predict the rainfall in India. India Meteorological Department always provides healthy prediction of rainfall in India but the climatic condition of Assam and Meghalaya affects the prediction of rainfall in Assam. Time

U. Barman (✉) · A. E. Hussain · M. J. Dahal · P. Barman · M. Hazarika
Department of Computer Science and Engineering, Girijananda Chowdhury Institute of Management and Technology, Guwahati, Assam, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_35

series analysis is already done by many researchers in the academic institute and research institute. In the last decade, researchers have done the forecasting of rainfall using different statistical and machine learning methods such as seasonal auto-regressive integrated moving average (SARIMA) [1–5] and auto-regressive integrated moving average (ARIMA) [6–11], auto-regressive moving average (ARMA) [12], and emerging fuzzy time series (FST) [9]. Among all the methods, SARIMA is one of the most popular techniques used by many researchers to forecast the rainfall time series. The paper [5] used the SARIMA model for the rainfall time series analysis of Shouguang city of China. The study [6] recorded ARIMA model for the rainfall analysis of Allahabad region of India. The paper [1] used SARIMA model for the time series analysis of Brong Ahafo Region of Ghana. Again, ARIMA and FST model are recorded [9] for the rainfall time series analysis of Nigeria. The Indian author [10] used ARIMA model for the rainfall time series analysis of Varanasi. The authors [12] used ARMA model for the forecasting of rainfall data in Colorado, USA.

From the above examples, it is found that the statistical methods are often used by many researchers in their research. Apart from these methods, time series forecasting is also done using different machines learning methods such as recurrent neural network (RNN) [13], support vector regression (SVR) [13–15], and artificial neural network (ANN) [16]. Among all the statistical and artificial intelligence-based methods, SARIMA is widely used in rainfall forecasting. It is because of the seasonality effect handling parameter of SARIMA. The SARIMA model is dependent only on time but not on the other climatic conditions such as temperature, pressure, and humidity. Since the intention is to forecast the Assam rainfall with time, the SARIMA model is used in this paper, and the result is compared with ARIMA.

35.2 Material and Methods

35.2.1 About Dataset

The dataset of the paper is the meteorological data. The dataset contains the monthly rainfall data from 1901 to 2017. The visualization of the rainfall data is presented in Fig. 35.1.

35.2.2 Preprocessing and Components of Rainfall Data

The ARIMA and SARIMA time series analysis is possible when the data has no trend, seasonality, and irregularity. But the trend and seasonality are the prime factors to be considered before applying the time series analysis. In this paper, we build our model based on the three components of the model. Let $y(t)$ be the analysis output of the model, then $y(t)$ can be express as below

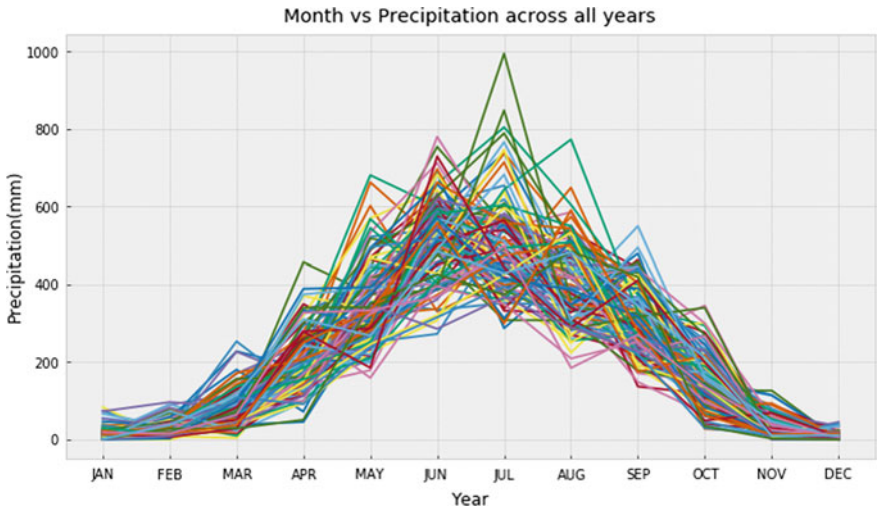


Fig. 35.1 Data plot of precipitation versus months

$$Y(t) = F(\text{Trend},\ \text{Seasonality},\ \text{Irregularity}) \tag{35.1}$$

To ensure the trend and seasonality, the rainfall time series data is divided into its components, and it is presented in Fig. 35.2.

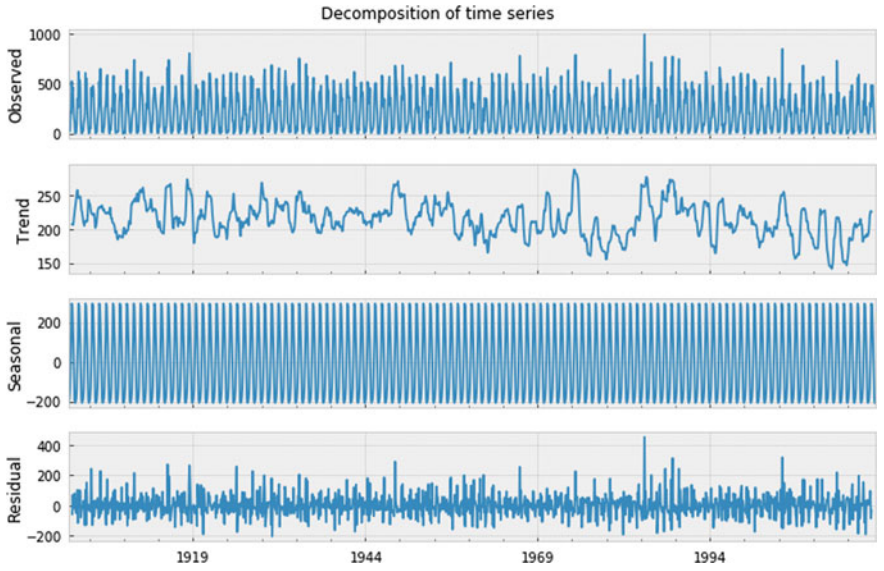


Fig. 35.2 Trend, seasonality, and residual of the rainfall data

Table 35.1 Result of ADF test

Parameter	Value
Test statistic	−5.299084
p-value	0.000005
#Lags used	24.00000
Number of observations used	1379.000000
Critical value (1%)	−3.435101
Critical value (5%)	−2.863638
Critical value (10%)	−2.567887

Figure 35.2 defines that the rainfall data does not have any trend and high value of residual. But the data contains seasonality as rainfall is seasonal in Assam, and it is more from April to November as compared to other months of the year. There is no meaning of apply time series analysis on the seasonal data. To valid the points, we have performed augmented Dickey–Fuller (ADF) test to check the stationary of the rainfall time series data. The ADF test is a statistical tool which is used to find the stationarity of the time series data. It calculates the unit root of the data based on which the stationarity is calculated. In this paper, two hypotheses are considered for ADF test. One is null, and another is alternate hypothesis. The null hypothesis is that the time series rainfall data has a unit root. The alternate hypothesis is that the rainfall series data has not any unit root. The rejection and acceptance of the null hypothesis are depended on the *p*-value of the ADF test. The ADF test is applied in python environment. The outcome of the ADF test is presented in Table 35.1.

In this paper, the *p*-value of the ADF test is 0.000005 which is very less as compared to the threshold value of the *P*. The threshold value of *p* is 0.05 or 0.01. Apart from the *p*-value, the test statistic is −5.299084 which is less than the critical value of the ADF test at 1%. These two points reject the null hypothesis and consider the data as stationary. The data is stationary means it is independent of time. To valid the point, the moving average (MA) or rolling mean is applied on the train data with different window size such as 4, 8, 12, and 16. The MA calculates the mean of the train data at specific time instant *t*. In this paper, we consider the time instant as window size. The output of the MA is represented in Fig. 35.3 using graph plots.

35.2.3 Time Series Analysis of Rainfall Data Using ARIMA and SARIMA

In this paper, the SARIMA and ARIMA models are applied for the time series analysis of rainfall time series data. SARIMA is the next version of ARIMA model. ARIMA model consists of auto-regression (AR), moving average (MA), and integration of AR and MA. The ARIMA model is denoted using three non-seasonal parameters such as *p*, *d*, and *q*. The parameter *p* is the order or the lag value of the AR model,

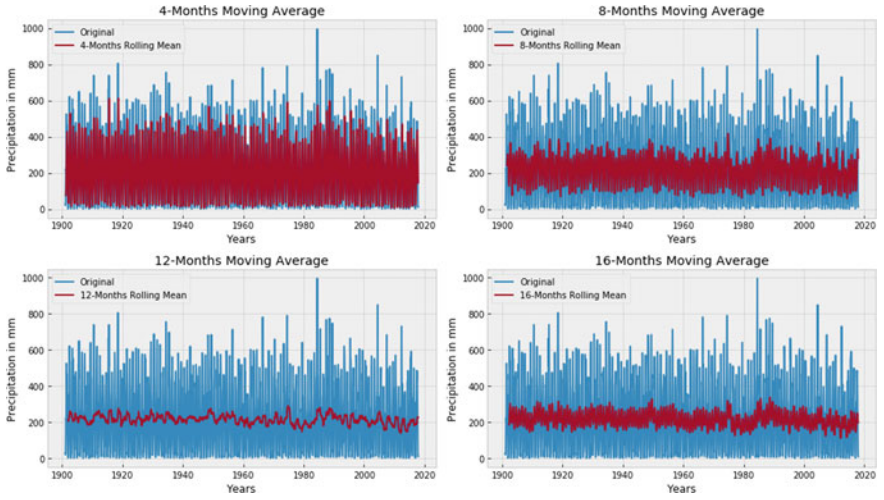


Fig. 35.3 Moving average with different window

whereas d and q are the degrees of integration and MA model. The SARIMA consists of two elements such as trend and seasonal. The trend component is defined by p, d, and q where p is the autoregressive order, q is the moving average order and d is the order of difference. The seasonal components consist of P, D, Q and m where P is the order of seasonal autoregressive, Q is the order of seasonal moving average order, D is the order of difference and m is the single period step. In this paper, the value of m is considered as 12. For the best value p, d, q, P, D, and Q, the different combinations of p, d, and q within the range 0–2 are calculated. These combinations are passed to the SARIMA as a parameter and the best combination is found based on the value of the AIC, BIC, and HQIC. The different combinations of trend and seasonal parameters are presented in Table 35.2.

Among all the combinations, the best combination is $p = 1$, $d = 0$, $q = 1$, $P = 0$, $D = 1$, $Q = 1$ and $m = 12$ with AIC = 12,846.66057.

Table 35.2 Trend and seasonal component of SARIMA

Trend combination	Seasonal combination
0, 0, 0	0, 0, 0, 12
0, 0, 1	0, 0, 1, 12
0, 1, 0	0, 1, 0, 12
0, 1, 1	0, 1, 1, 12
1, 0, 0	1, 0, 0, 12
1, 1, 0	1, 1, 0, 12
1, 1, 1	1, 1, 1, 12

35.3 Results and Discussion

Along with the SARIMA model, we have also applied the ARIMA model on the Assam rainfall time series data. Since the initial rainfall data shows that the time series doesn't have any seasonal effect, so the ARIMA model is applied on the time series data. The summary of the ARIMA and SARIMA model are shown in Table 35.3. The paper [11] shows that the AIC and BIC are the efficient parameters to select the best SARIMA and ARIMA model for time series. A model fits well when the value of AIC and BIC is minimum [11]. In this paper, the value of AIC is more in ARIMA (177,761.970) as compared to SARIMA (12,846.661) (Table 35.3).

The values BIC and HQIC of SARIMA are 12,866.731 and 12,854.248, respectively. But the values of BIC and HQIC for ARIMA are 17,782.959 and 17,769.81, respectively. Along with AIC, BIC, and HQIC, values of SARIMA have proved that SARIMA is better than ARIMA for the seasonal rainfall data. It means that ARIMA model is not an efficient techniques to predict the future rainfall in Assam. The SARIMA model shows the better result as compared to the ARIMA model due to the seasonality effect of the rainfall time series data. Along with AIC, BIC, and HQIC, the two models are evaluated using the regression score (RC), mean absolute error (MAE), median absolute error (MeAE), mean squared error (MSE), mean squared log error (MSLE), and root mean square error (RMSE). The evaluating parameters of the models are shown in Table 35.4.

In the previous study, the authors used coefficients of regression ($R2 = 0.86$) [11], RMSE (18.60 for 1 days forecast) [14], MeAE (0.3426) [15], and MAE (15.26) [7] for evaluating the different time series analysis model. Here, both the models are compared based on the mentioned evaluating parameters. The RMSE value of SARIMA is 84.40 whereas the RMSE of ARIMA is 135.71. The RC, MAE, MeAE, MSE, and MSLE of SARIMA are 0.79, 55.86, 30.42, 7124.05, and 0.4171, respectively, whereas the RC, MAE, MeAE, MSE, and MSLE of ARIMA is 0.48884, 101.57, 69.57, 18,417.62, and 2.0587, respectively. The evaluating parameters clarify that the SARIMA is much more efficient as compared to ARIMA.

Table 35.3 SARIMA and ARIMA model summary

Model	Train sample	AIC	BIC	HQIC
SARIMA	1901:1994	12,846.661	12,866.731	12,854.248
ARIMA	1901:1994	17,761.970	17,782.959	17,769.81

Table 35.4 Evaluating parameters of SARIMA and ARIMA

Model	RC	MAE	MeAE	MSE	MSLE	RMSE
SARIMA	0.79	55.86	30.42	7124.05	0.4171	84.40
ARIMA	0.4884	101.57	69.57	18,417.64	2.0587	135.71

35.4 Conclusion

This paper is to study the rainfall forecast in Assam State only, and the SARIMA model is found as the best suitable method to carry out the research. In this paper, both the ARIMA and SARIMA models are applied for the rainfall prediction in Assam. The SARIMA model is found to be accurate and efficient as compared to ARIMA. SARIMA model is helpful as the dataset contains the seasonal effect in the data. The SARIMA is helpful to forecast the rainfall in future. Time is the key parameter to study about the rainfall forecast, and it can be used to predict the rainfall in long-range forecasting using SARIMA. In future, the Assam rainfall can be analyzed with help of SVR, ANN, and KNN regression, and the result can be compared with the ARIMA and SARIMA.

Acknowledgements This research project is supported by Assam Science and Technology University, Guwahati, Assam under TEQIP-III, vide Ref. No.: ASTU/TEQIP-III/Collaborative Research/2019/2474, Dated July 17, 2019.

References

1. Afrifa-Yamoah, E., Saeed, B.I., Karim, A.: Sarima modelling and forecasting of monthly rainfall in the Brong Ahafo Region of Ghana. World Environ. **6**(1), 1–9 (2016)
2. Dabral, P.P., Murry, M.Z.: Modelling and forecasting of rainfall time series using SARIMA. Environ. Process. **4**(2), 399–419 (2017)
3. Eni, D., Adeyeye, F.J.: Seasonal ARIMA modeling and forecasting of rainfall in Warri Town, Nigeria. J. Geosci. Environ. Prot. **3**(06), 91 (2015)
4. Subbaiah Naidu, K.: SARIMA modeling and forecasting of seasonal rainfall patterns in India. Int. J. Math. Trends Technol. (IJMTT) **38**(1), 15–22 (2016)
5. Wang, S., Feng, J., Liu, G.: Application of seasonal time series model in the precipitation forecast. Math. Comput. Model. **58**(3–4), 677–683 (2013)
6. Graham, A., Mishra, E.P.: Time series analysis model to forecast rainfall for Allahabad region. J. Pharmacogn. Phytochem. **6**(5), 1418–1421 (2017)
7. Mohamed, T.M.: Time series analysis of Nyala rainfall using ARIMA method (2016)
8. Murat, M., Malinowska, I., Gos, M., Krzyszczak, J.: Forecasting daily meteorological time series using ARIMA and regression models. Int. Agrophys. **32**(2), 253–264 (2018)
9. Olatayo, T.O., Taiwo, A.I.: Statistical modelling and prediction of rainfall time series data. Glob. J. Comput. Sci. Technol. (2014)
10. Shivhare, N., Rahul, A.K., Dwivedi, S.B., Dikshit, P.K.S.: ARIMA based daily weather forecasting tool: a case study for Varanasi. Mausam **70**(1), 133–140 (2019)
11. Swain, S., Nandi, S., Patel, P.: Development of an ARIMA model for monthly rainfall forecasting over Khordha district, Odisha, India. In: Recent Findings in Intelligent Computing Techniques, pp. 325–331. Springer (2018)
12. Burlando, P., Rosso, R., Cadavid, L.G., Salas, J.D.: Forecasting of short-term rainfall using ARMA models. J. Hydrol. **144**(1–4), 193–211 (1993)
13. Hong, W.-C.: Rainfall forecasting by technological machine learning models. Appl. Math. Comput. **200**(1), 41–57 (2008)
14. Hasan, N., Nath, N.C., Rasel, R.I.: A support vector regression model for forecasting rainfall. In: 2015 2nd International Conference on Electrical Information and Communication Technologies (EICT), pp. 554–559 (2015)

15. Zhao, S., Wang, L.: The model of rainfall forecasting by support vector regression based on particle swarm optimization algorithms. In: Life System Modeling and Intelligent Computing, pp. 110–119. Springer (2010)
16. Abhishek, K., Singh, M.P., Ghosh, S., Anand, A.: Weather forecasting model using artificial neural network. Procedia Technol. **4**, 311–318 (2012)

Chapter 36
Reliable Low-Cost High-Speed and Direction-Controlled Surveillance Robot



S. Rooban, I. Venkata Sai Eshwar, R. Radhika, M. Dheeraj Kumar, and K. Chaitanya Krishna

Abstract Reliability is the basic requirement for all the products and even for human beings. All the researchers and manufacturing industry are focusing toward improving the reliability of the products or circuits with minimum cost. This paper presents a speed and direction control of cost-effective robot which can be used for surveillance in places where human beings are unable to enter. Many surveillance robots have been proposed by many researchers in recent years. Always the customer expects single device usage for various applications; this requirement makes the designer to incorporate all the features in a device that leads to increased cost and reduction in speed. One of the methods to improve the speed of the device processing capability with all features is parallel processing that makes the product costlier because high cost of high-speed parallel processor. The proposed method incorporates many features and increases the speed of operation by doing parallel operation with different processors. This proposed method of designing a surveillance robot uses three processors which are Arduino, MyRIO and NodeMCU. All the three processors are of low cost and are easily available. The proposed method is divided into modules, and various features of surveillance robot are allocated into each module with different processors. So that robot can operate in parallel for the applications. Because of the parallel operation, the speed of operation is improved, and if any one processor fails, the robot can work with the limited features whereas in single processor once the processor fails that collapse the robot.

S. Rooban (✉) · I. V. S. Eshwar · M. D. Kumar · K. C. Krishna
Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh 522502, India

R. Radhika
St. Peter's College of Engineering and Technology, Avadi, Tamil Nadu 600054, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_36

36.1 Introduction

With increasing popularity of robotics and automation in industry 4.0, the need for reliability is increased and becomes a serious challenge. Surveillance is a necessary requirement in all the industry and public places. With the increasing usage of machine learning [1] and artificial intelligence techniques, the maximum input for those techniques is through sensors. Many applications require continues data from the places where humans cannot be roaming all the time to collect the data. Robot can be used everywhere to sense or collect the information and passed to the system which needs the information. How fast the robot completes the task or responds for the instruction is also one of the major requirements in the industry 4.0. Speed and reliability play a major role in choosing the robot to meet our required application from the various robot manufacturers. Processor which is used for the robot provides more contribution to the speed and reliability. There are many research works in robot such as path planning [2], algorithmic [3] and mathematical [4] are proposed earlier by many researchers to improve the reliability. The proposed method is one of the simple solutions to improve the reliability by keeping the cost of the robotics system as same. The remaining paper is presented as follows. Section 36.2 describes the work related to surveilling robots. Section 36.3 presents the proposed method with various processors. Section 36.4 discusses the results and the paper is concluded in Sect. 36.5.

36.2 Related Work

In the past years, it is observed that improvement in reliability of any system is the basic requirement. Reliability can be improved by improving the processor internally such as the adder [5] also by adopting some architectural modifications [6]. Various processors [7] are analyzed and conclusion arrived that the reliability also depends on performance of processor [8, 9]. Robot performance can be improved by efficient path planning [10] and the hardware components especially the sensors [11]. Many methods are proposed by researchers to improve the reliability by using reliable sensors [12, 13]. Techniques [14] are adopted to improve the reliability of the processor in chip manufacturing and testing phase. Parallel processing can make the processor to perform the task in high speed. In this work, processing of all the features of the robot is performed in various processors to improve the reliability and speed.

36.3 Proposed Methodology

The proposed surveillance robot features are divided into three modules, they are (1) patrolling unit, (2) vision unit and (3) environmental identification unit. As we have discussed earlier the parallel processing is obtained by using three simple cost-effective processors for each module. Patrolling unit uses Arduino, vision unit uses MyRIO, and environmental unit uses NodeMCU.

36.3.1 Patrolling Unit

The block diagram for the patrolling unit is shown in Fig. 36.1. As our work is distinguished from others by low cost, the hardware parts such as DC motor, Motor driver, Arduino, Bluetooth module and mobile are easily available with minimum cost. In surveillance system, patrolling is an important module. This module is responsible for moving the robot structure into any places where humans are hazardous to the environment or the environment is hazardous to humans. The movement is obtained

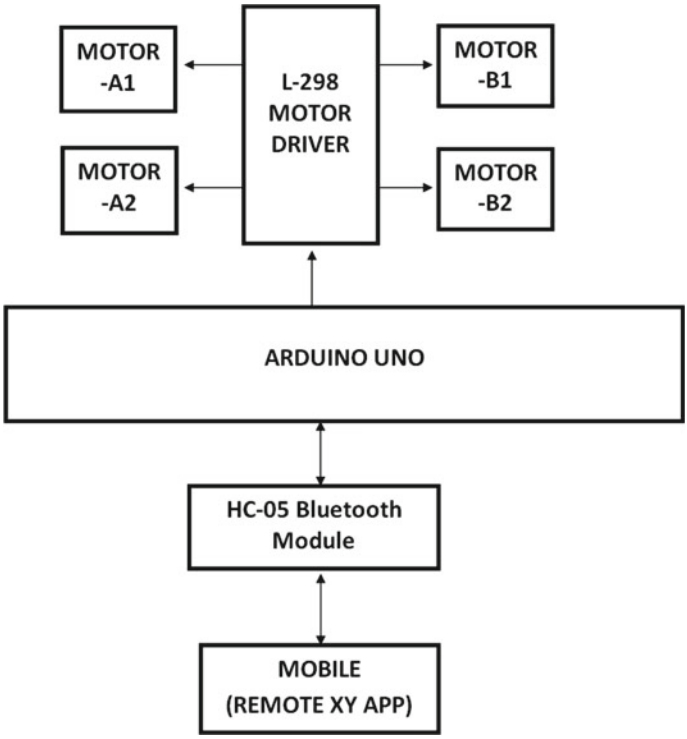


Fig. 36.1 Block diagram for patrolling unit

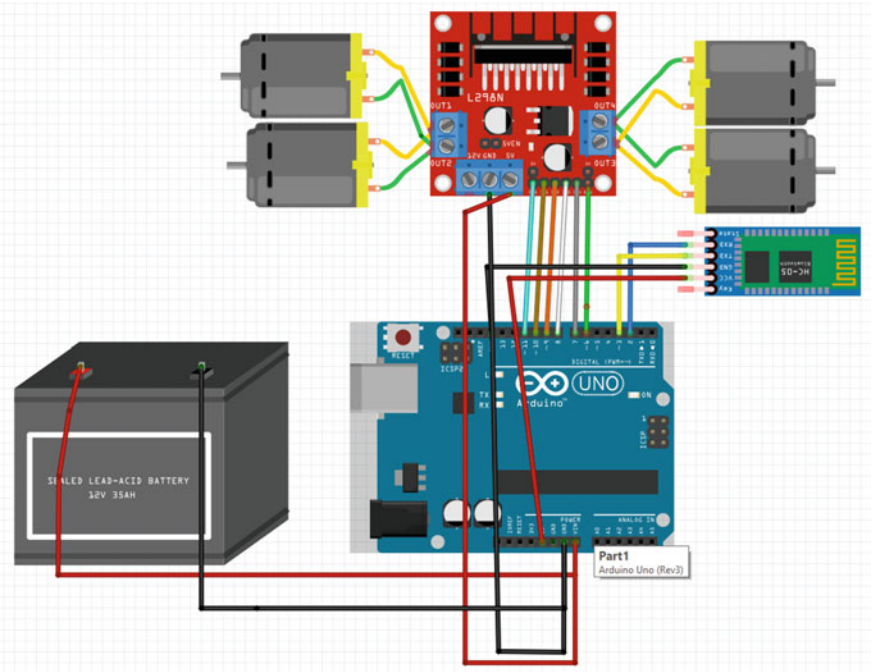


Fig. 36.2 Circuit connection diagram of patrolling unit

by the four DC motors. Controlling of the robot can be done using communication between Bluetooth module and mobile which is installed with the simple Remote XY Arduino application.

Connection diagram of the patrolling unit is shown in Fig. 36.2. The direction of movement for the patrolling depends on the activation of the respective DC motors. For example, if the robot is required to move straight, all the motors have to be activated in forward direction, and if the robot is required to turn left, the two right motors will be activated. Also the speed is controlled through the number of pulses applied to the motor per second. Pulse width modulation technique is used for the number of pulses to be applied to meet the required speed. Activation of the motor based on the number of pulses applied for the required patrolling is written as instruction in any one of the programming language python, C, C++ are dumped into the Arduino.

Figure 36.3 shows the Remote XY application in the mobile. It is having the joystick icon, by using the joystick Icon in the proper movement the required instruction will be communicated to the Arduino through the Bluetooth device. The code takes the signal as input and executes the respective portions of the program to provide the required patrolling. It's a simple cost-efficient method of controlling the surveillance structure accurately.

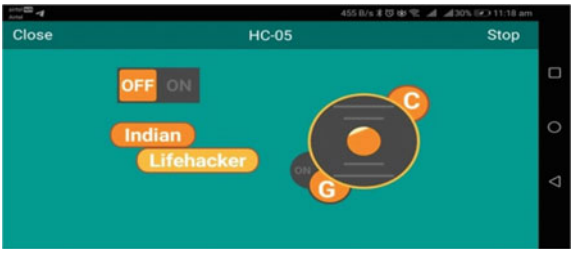


Fig. 36.3 Remote XY application

36.3.2 Vision Unit

As our proposed work is focused on dividing the features into modules and run the modules in separate simple processors to make use of good performance with the parallel processing [15]. Vision of a surveillance robot is an important input to the system. To make this module as simple as possible no complexity is added in the vision unit to achieve the lowest cost for the system. MyRIO processor is used with the webcam and the live telecast can be viewed in the laptop. This module can be extended by using Bluetooth to the MyRIO to transfer the video to the systems to record the videos in the laptop by installing the video recording. MyRIO with the LabVIEW can control the Robot independently with better performance by using the various toolbox in the LabVIEW robot toolbox. Figure 36.4 shows the flow diagram, Figs. 36.5 and 36.6 show the simple implementation of the MyRIO with LabVIEW and webcam.



Fig. 36.4 Flow diagram of vision unit



Fig. 36.5 Simple MyRIO setup

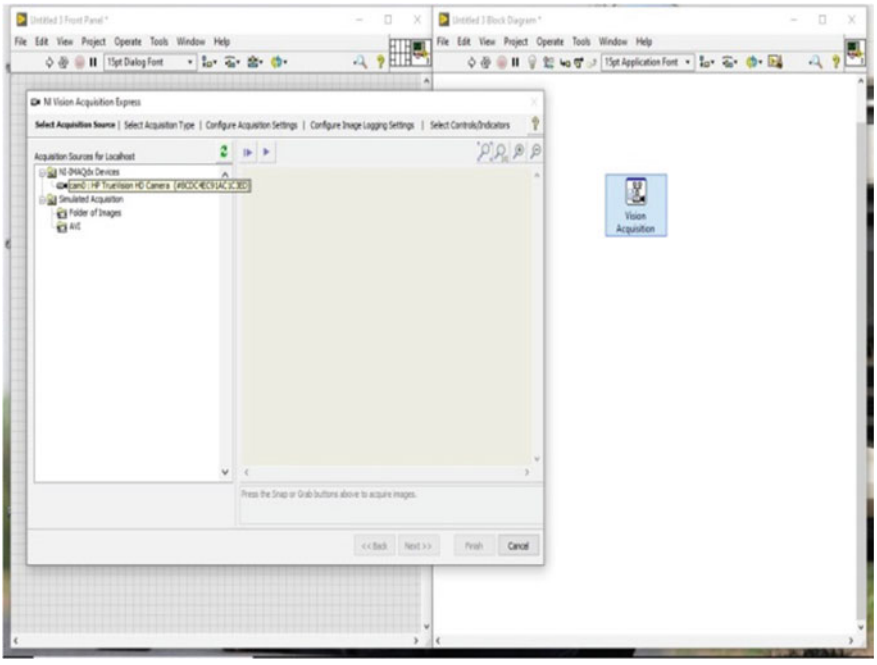


Fig. 36.6 LabVIEW with the vision unit

36.3.3 Environmental Identification Unit

Most important features and the basic requirement of surveillance robot is to collect the environment status through various sensors and communicate the collected information to the processor then the processed information is transferred and delivered in useful form. A simple setup of all the sensors and the processor is shown in Fig. 36.7. NodeMCU is used as processor, DHT11 sensor for humidity and temperature sensing, Mq2 sensor for sensing the Gas [12]. Servo motor is attached to the system to make any mechanical movement by the processor based on the requirement.

Figure 36.8 shows the connection diagram of the Environmental Identification Unit. This module is controlled with the installed blink application in mobile. Bluetooth module communicates the status of various sensors immediately to the mobile through the blynk. This is most useful for keep track of any environmental variations continuously. Figure 36.9 shows the output of the sensors which is used in Fig. 36.7.

36.4 Results

The proposed reliable high-speed and direction-controlled surveillance robot is completed with the three processors with various blocks. Figure 36.10 shows the hardware of the patrolling module in the system. Figure 36.11 shows the output of MyRIO [7] with the LabVIEW. Figure 36.12 shows the final structure of the proposed surveillance system. It is tested with all the possible inputs, and the systems work well. Speed of the robot is checked by varying the instruction through the joystick icon in the mobile.

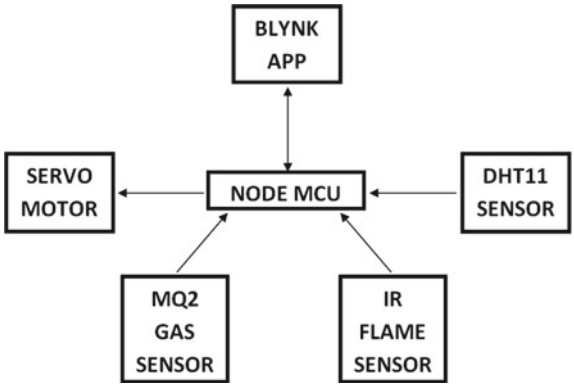


Fig. 36.7 Block diagram of environmental identification unit

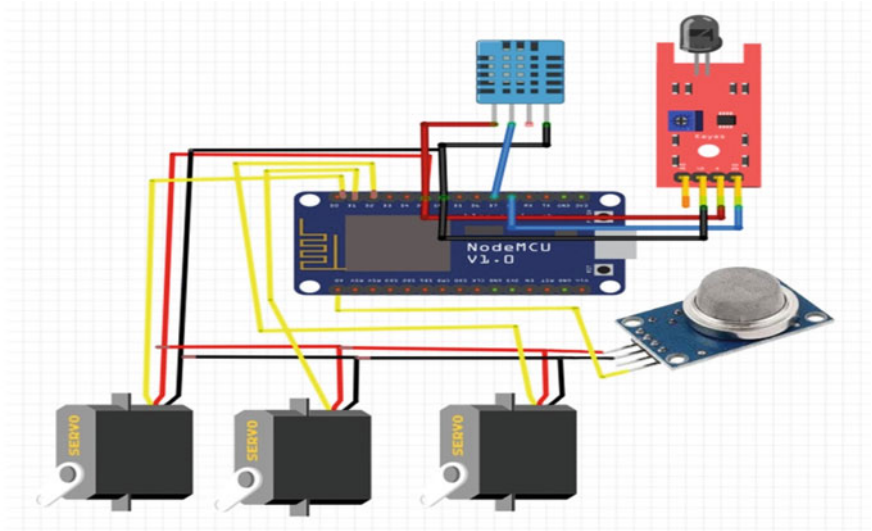
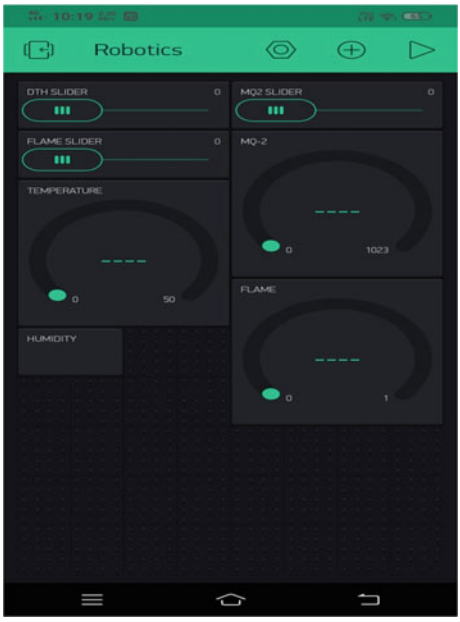


Fig. 36.8 Circuit connection diagram of environmental identification unit

Fig. 36.9 Output of sensors at blynk application



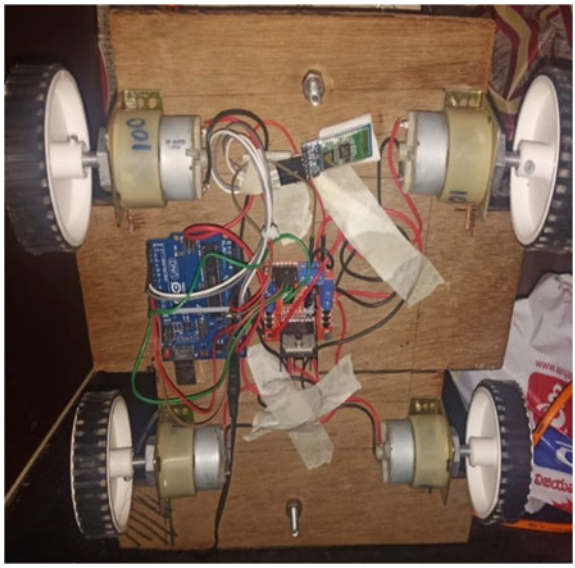


Fig. 36.10 Patrolling hardware

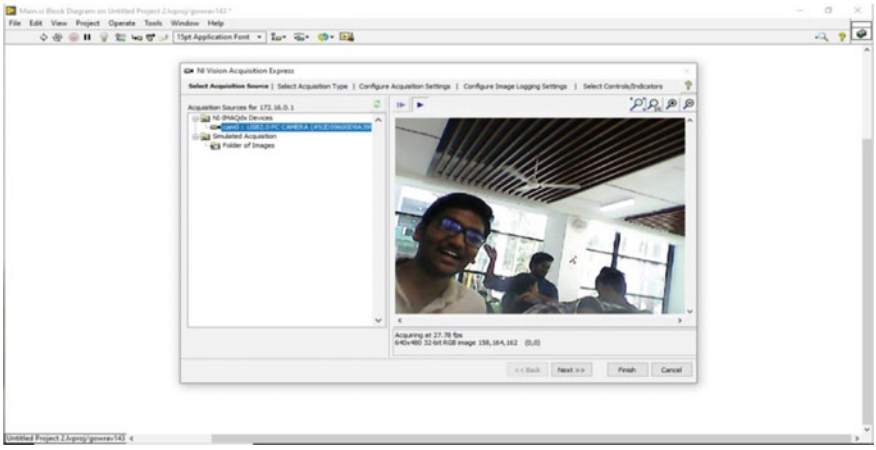


Fig. 36.11 MyRIO output in LabVIEW

36.5 Conclusion

This proposed work proves that reliability of the robot can be improved with minimum cost by sharing the required features of the robot with several low-cost processors. If any processor fails, robot can function with minimum features by using the available processors. Arduino, MyRIO and NodeMCU processors are used instead of going for the three same processors. Usage of different processor can also improve the reliability because of specification of each processor differs which gives



Fig. 36.12 Final proposed reliable surveillance system

assurance that all the processor never fails at the same time. Future extension of this work is that to develop a communication or interconnection between all the sensors and motors so that if one processor fails, the corresponding sensors and motors can be activated by the other processors. This will make the system more reliable without sacrificing any features.

References

1. Danthala, S., Rao, S., Mannepalli, K., Shilpa, D.: Robotic manipulator control by using machine learning algorithms: a review. Int. J. Mech. Prod. Eng. Res. Dev. **8**(5), 305–310 (2018)
2. Victerpaul, P., Saravanan, D., Janakiraman, S., Pradeep, J.: Path planning of autonomous mobile robots: a survey and comparison. J. Adv. Res. Dyn. Control Syst. **9**(Special Issue 12), 1535–1565 (2017)
3. Gopi Krishna, P., Sreenivasa Ravi, K., Bhavya Sai Sireesha, P.: Implementation of bi-directional bluetooth—ZigBee gateway for multipurpose applications in iot. J. Adv. Res. Dyn. Control Syst. **9**(Special Issue 18), 306–317 (2017)
4. Vani, G.D., Chinnaiah, M., Karumuri, S.R.: Hardware scheme for autonomous docking algorithm using FPGA based mobile robot. In: Proceedings of the 2018 8th International Symposium on Embedded Computing and System Design, ISED 2018, pp. 110–115 (2018)
5. Venkateswara Rao, M., Madhav, B.T.P., Anilkumar, T., Prudhvi Nadh, B.: Metamaterial inspired quad band circularly polarized antenna for WLAN/ISM/Bluetooth/WiMAX and satellite communication applications. AEU Int. J. Electron. Commun. **97**, 229–241 (2018)
6. Rooban, S., Saifuddin, S., Leelamadhuri, S., Waajeed, S.: Design of fir filter using wallace tree multiplier with kogge-stone adder. Int. J. Innov. Technol. Explor. Eng. **8**(6), 92–96 (2019)
7. Naveen, P.S., Rao, K.R.R.M.: Ni myrio based smart robot with authentication switch. Int. J. Innov. Technol. Explor. Eng. **8**(6), 88–91 (2019)

8. Anitha, A., Rooban, S., Sujatha, M.: Implementation of energy efficient gates using adiabatic logic for low power applications. Int. J. Recent Technol. Eng. **8**(3), 3327–3332 (2019)
9. Rooban, S., Kumar, K.S., Shankar, K.R., Bhaskara Rao, N.U.: Test and analysis of high performance microprocessor through power binning method. Int. J. Eng. Adv. Technol. **8**(4), 882–886 (2019)
10. Pallav, K., Siva Prasanth, G.V.S., Shanmukha Sree Charan, T.P.B.M., Nayak, D.K., Selvakumar, R.: Design and simulation of path planning algorithm for autonomous mobile robot navigation system using ekfslam. Int. J. Sci. Technol. Res. **8**(12), 1333–1339 (2019)
11. Prasad, C.H., Bojja, P.: A reliable, energy aware and stable topology for bio-sensors in healthcare applications. J. Commun. **14**(5), 390–395 (2019)
12. Sunithamani, S., Sanjay Naidu, R., Sudheep, V., Hemanthkumar, C.: Analysis of MEMS based gas sensor. Test Eng. Manag. **81**(44147), 6186–6189 (2019)
13. Rajasekar, K., Sunithamani, S.: Low setting time offering series capacitive RF MEMS switch for WI-FI applications. Int. J. Recent Technol. Eng. **7**(5), 439–442 (2019)
14. Soumya, N., Sai Kumar, K., Raghava Rao, K., Rooban, S., Sampath Kuma, R.P., Santhosh Kumar, G.N.: 4-bit multiplier design using cmos gates in electric VLSI. Int. J. Recent Technol. Eng. **8**(2), 1172–1177 (2019)
15. Radhika, R., Partheeban, N.: Improving power efficiency in the implementation of discrete cosine transform in multicore architecture. J. Comput. Theor. Nanosci. **15**(8), 2597–2603 (2019)

Chapter 37
Deciphering WEP, WPA, and WPA2 Pre-shared Keys Using Fluxion



Sidharth Atluri and Revanth Rallabandi

Abstract The content in the present paper reviews the attack carried out to decipher WEP, WPA, and WPA2 passwords using a tool named Fluxion. Before understanding how the tool works, brief information about the security protocols is given so that comprehending the attack more undoubtedly. At the same time, we compare Fluxion to other pre-existing tools used to crack Wi-Fi passwords and then list out the major differences between them. This comparison is done to evaluate the efficacy and success rate of Fluxion. In conclusion, this research and attack are performed to provide a deeper insight into WEP, WPA, and WPA2 cracking.

37.1 Introduction

37.1.1 WEP

One of the oldest security protocols is the Wired Equivalent Privacy or WEP. WEP is not being used in the present due to its vulnerabilities and the fact that it can be easily cracked. The underlying algorithm behind this security protocol is RC4. The way this algorithm works is, whenever a client wants to send data to the router or vice-versa, first the data will be encrypted using a key. Therefore, this will convert the plain text in the packet to ciphertext. This encrypted packet is now sent in the air. The recipient will be able to transform the packet back to its original form because it has the key. The problem arises with the way WEP implements this algorithm. To encrypt the packet WEP generates a unique keystream. In other words, for each packet that is being sent into the air WEP tries to create a keystream for it. The keystream is a combination of a random 24-bit initialization vector (IV) added to the pre-shared key (password) of the network. Before sending this encrypted packet in

S. Atluri (✉)
Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

R. Rallabandi
Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_37

the air, WEP appends the 24-bit IV to the encrypted packet to ensure that the recipient can decrypt the packet because it needs both the IV and the pre-shared key. Now the major issue here is that the IV that is being appended is in plain text. Also, the size of an IV is only 24 bits. When we compare it to the amount of traffic generated on a Wi-Fi network, it is very less. Consequently, IVs will start getting repeated.

37.1.2 WPA/WPA2

With WEP not meeting the specific security requirements, Wi-Fi Protected Access (WPA) came into the picture in the year 2003. WPA provided a more complex data encryption procedure and it also had a finer user authentication when compared to WEP. In 2004, WPA2 was created which had some major upgrades from WPA. To begin with, it must be understood that both are very similar. The difference arises in the encryption method used to maintain message integrity. WEP uses Temporal Key Integrity Protocol (TKIP) whereas WPA2 makes use of Advanced Encryption Standard (AES). But this does not affect the methods used to crack either WPA or WPA2 networks. WPA/WPA2 came over the weaknesses in WEP by not appending any useful data in the packets to be sent in the air. In addition to that, the keys are unique, temporary, and exist for a longer period than WEP. Hence, the only packets that help in cracking this network are the four handshake packets that are exchanged between the client and the router when the client connects to it.

37.2 Related Work

In [1], the author tests all the router protocols available currently such as WEP, WPA, WPA2, and WPS and detects the vulnerabilities of the system. The author also explains the tools like Airodump-ng, Aircrack-ng that is used to acquire access point pins. The author in [2] uses the Aircrack-ng tool for cracking the WPA pre-shared key with simple ASCII keys and a complex hexa-decimal key to check if keys could be cracked. Kissi and Asante [3] discuss the usage of penetration testing to assess vulnerabilities and conduct attacks on WEP, WPA and WPA2 security protocols. In [4], the author concludes by comparing Fluxion with previously defined hypotheses and aims to give more insight toward the WEP/WPA/WPA2 attacks.

37.3 Existing System

The very first step is to configure the wireless interface connected to monitor mode. Changing the mode of the wireless adapter can be done in the manner shown below. Here, we can notice the status of the wireless adapter ‘wlan0’ by using the command

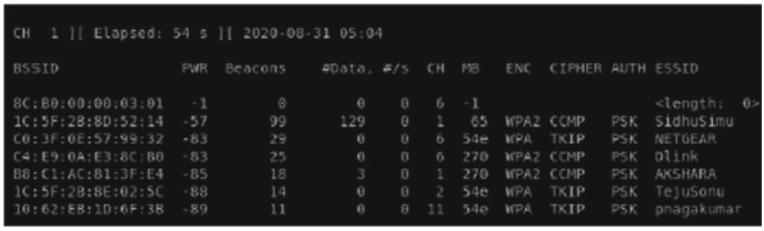


Fig. 37.1 List of networks in Airodump-ng

'iwconfig'. When we look at the mode attribute, it has been changed from managed to monitor. Now the adapter is ready to capture any packets being sent in the air. The next step is to find out the available networks and select the network on which we want to perform the attack (Fig. 37.1).

By running Airodump-ng we can obtain the list of networks as shown above. For example, the attack will be carried out on the wireless network named 'SidhuSimu' which has WPA2 protocol enabled.

The next step is to capture the handshake by using the following command as shown in Fig. 37.2.

The catch is that a handshake can only be captured when a new client connects to the network. One way is to wait for a new client to connect to the network. Alternatively, we can use a de-authentication attack, by which we can disconnect a client from the network for a short period. The client automatically reconnects to the network once we stop the attack. Therefore, the handshake will be sent when the client automatically reconnects, consequently allowing us to not wait for an actual connection to take place (Fig. 37.3).

The final step is to use Aircrack-ng and a wordlist to crack the pre-shared key from the handshake as shown in Fig. 37.4.

Therefore, the pre-shared key for the above network was found. This is the typical way of cracking WPA/WPA2 networks and even WEP networks.



Fig. 37.2 Command used to capture the handshake

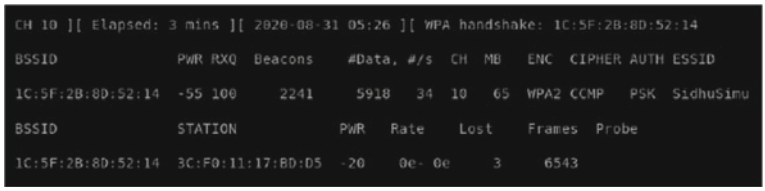


Fig. 37.3 Success page indicating the immediate capture of handshake



Fig. 37.4 Loading the word list used to crack the PSK in Airodump-ng

37.4 Proposed System

Even though the above-depicted method works in most cases, it might be difficult to crack the password if the wordlist generated does not contain the pre-shared key. In addition to that, capturing handshakes, performing de-authentication attacks, and manually typing the commands are time-consuming. Here, we can make use of Fluxion which provides us with a wide variety of attacks and zero requirements of running commands manually. Fluxion can be considered as a mixture of social engineering and technical automation that fools the user into giving the Wi-Fi passwords in no time. To be concise, it is more of a social engineering framework that makes use of an evil twin access point, handshake capture functions, and integrated jamming.

Fluxion is a distinctive tool in its utilization of a WPA handshake to not only restricts the functioning of the login page, but also the functioning of the entire script. It deactivates the original network and establishes a duplicate with the same name, persuading the disconnected client to join. It then displays a fake login page mentioning the router requires restarting to install firmware and it needs the password to do so. The tool makes use of a captured handshake to verify the pre-shared key submitted and continues to deactivate the target access point (AP) until the right password is typed.

37.4.1 Installation

The following command as shown in Fig. 37.5 installs Fluxion.

We need to check if there are any missing dependencies by moving to the 'Fluxion' folder and then running the Fluxion script as shown in Fig. 37.6 and 37.7.

Running the same command with –i option will automatically install the missing dependencies.



Fig. 37.5 Cloning the Fluxion git repository to install Fluxion



Fig. 37.6 Checking if there are any missing dependencies



Fig. 37.7 List of networks being displayed in Fluxion

37.4.2 Scanning for Wi-Fi Networks

After selecting the language, the tool provides us with a screen to select the kind of attack we want to perform as shown in Fig. 37.8.

The next screen provides us with a list of available wireless adapters that can be used to scan for networks as shown in Fig. 37.9.

Once the resource has been allocated, we move to a screen where we must select the frequency band in which we want to search for networks as shown in Fig. 37.10.

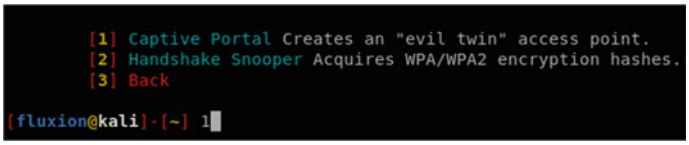


Fig. 37.8 List of attacks that can be performed

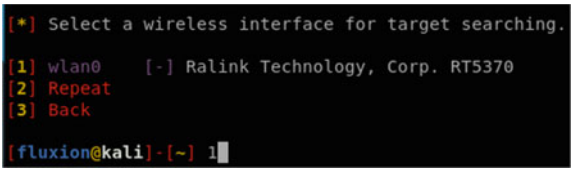


Fig. 37.9 List of wireless adapters

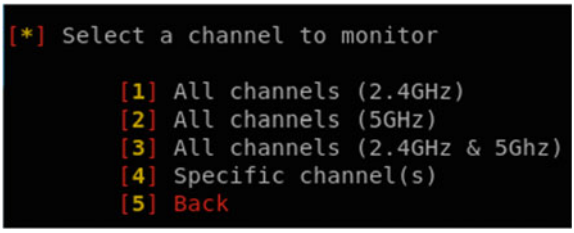


Fig. 37.10 List of available frequency bands

37.4.3 Selecting Target Access Point

We get a list of networks with other information such as encryption protocol being used, MAC address of the router, etc. Select a network from the given list as shown in Fig. 37.11. The next screen asks us to select a wireless adapter again. After selecting the wireless adapters, we need to select a method of de-authentication attack as shown in Fig. 37.12.

We also get a list of methods to implement fake AP service as shown in Fig. 37.13.

We are now given a list of password verification methods and hash verification methods as shown in Figs. 37.14 and 37.15.

Next, we must select the SSL certificate for the fake login page as shown in Fig. 37.16.

We get a list of fake login pages, and we can select one that might most probably trick the user as shown in Fig. 37.17.



Fig. 37.11 List of available Wi-Fi networks

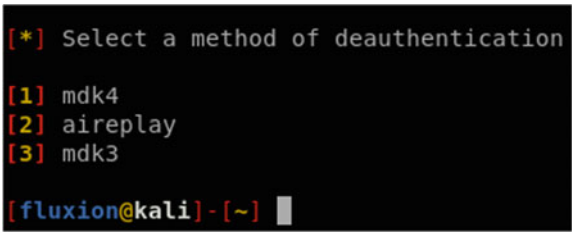


Fig. 37.12 List of methods for de-authentication attacks

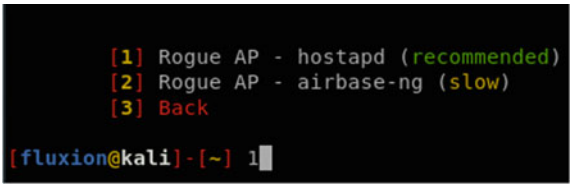


Fig. 37.13 List of fake access point services



Fig. 37.14 List of password verification methods

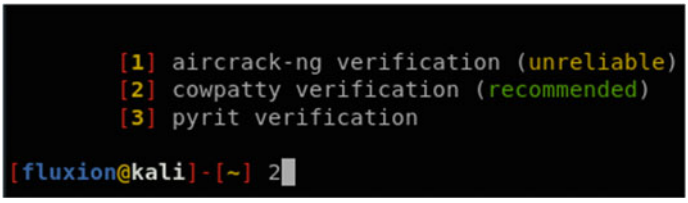


Fig. 37.15 List of hash verification methods

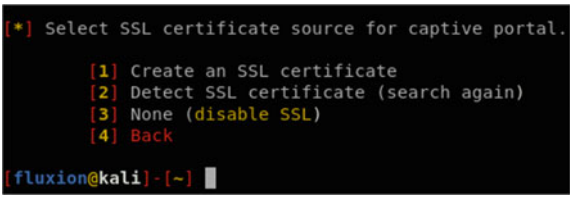


Fig. 37.16 List of fake SSL certificates



Fig. 37.17 List of fake login pages according to the router

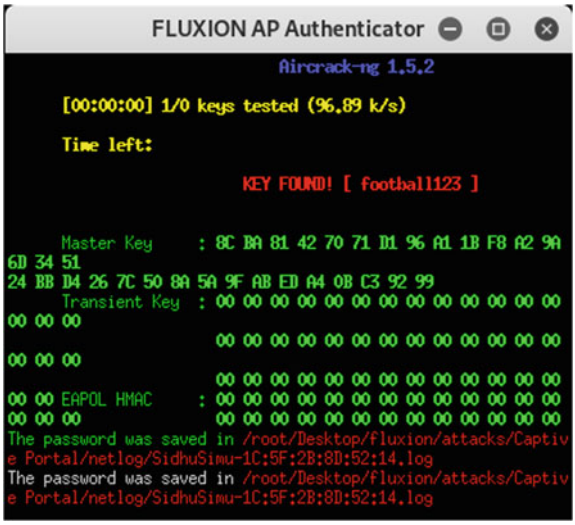


Fig. 37.18 Screen showing the hacked password of the Wi-Fi network

37.5 Results

The cracked password of the selected network 'SidhuSimu' and the directory in which the file is saved is shown in Fig. 37.18. Here, we can notice the efficiency of using Fluxion as a tool to crack passwords. Henceforth, we can also conclude that any security protocol that is in the market right now is still vulnerable to hackers and malicious attacks. There can be several other changes made to Fluxion 2.0 to carry out such attacks effortlessly.

37.6 Conclusion

The research portrays the advantages of using Fluxion over Aircrack-ng while cracking WEP/WPA/WPA2 networks. Unlike Aircrack-ng, the hacker does not need to memorize the commands to run de-authentication attacks, network jamming attacks, fake access point creation, handshake capturing etc. while using Fluxion as it provides a well-built user interface when compared to other cracking tools. The only limitation while using Fluxion is, if the target is an experienced IT professional or someone who understands computer networks, then they can figure out that it is a social engineering attack. In this case, the target might not share the pre-shared key (PSK) of the network. This paper signifies the importance of using proper security countermeasures because any user with minimal computer knowledge can exploit the vulnerabilities of a network, by simply learning the usage of a few security tools.

References

1. Pimple, N., Salunke, T., Pawar, U., Sangoi, J.: Wireless security—an approach towards secured Wi-Fi. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (2020)
2. Rabiul Islam, S.Md.: Wi-Fi protected access (WPA)—PSK (phase shift keying) key cracking using AIRCRACK-NG. Int. J. Sci. Eng. Res. (2013)
3. Kissi, M.K., Asante, M.: Penetration testing of IEEE 802.11 encryption protocols using Kali Linux hacking tools. Int. J. Comput. Appl.
4. Permatasari, F., Eaganathan, U.: An implementation of WEP/WPA/WPA2 password cracking using fluxion. In: 12th International Conference on Recent Innovations in Science, Engineering and Management (2018)

Chapter 38
Blockchain-Based Government Project Fund Tracking and Management System



M. V. Varada, Wissam Salih Abdulla, Aparna Udayakumar, Haritha Raghu, and Manoj V. Thomas

Abstract Across the globe, governments are testing blockchain in various administrative processes including public relations. Blockchain technology provides a secure platform and ensures data integrity. And its applications are resistant to breakdown and are secured using cryptography; hence, it can be used for government applications. Hyperledger fabric is one of the prominent open-source framework of blockchain which is used to develop a "permissioned" network system. When various project funds are issued by the government, often a large amount is unutilized for the actual project, due to corruption. It is necessary to leverage the hyperledger fabric for the formation of a network which allows efficient fund distribution systems resulting in a better society. In this paper, we explored adding permissioned support of hyperledger fabric to government fund distribution system.

38.1 Introduction

Corruption is a symptom and outcome of dishonesty or criminal offense undertaken by an institution. It has negative impact on the countries' GDP.

In the beginning of the novel technology, blockchain was used in cryptocurrency transactions. Presently, different other domains like health care and government started adopting this due to its decentralized and immutable nature [1]. Blockchain is now one of the popular technology which is at the center of the technology discussion. It is studied in research areas and applied in many fields, such as health care, education and business markets. In this paper, we focus mainly about the use of blockchain in a government domain. Specifically, this paper emphasizes the use of hyperledger fabric blockchain technology for tracking the funds generated by the government. Sometimes when funds are allocated to projects by the government, a large portion of it is unutilized, and it goes to the hands of various people involved in the project or some leaders, leading to corruption. So, if blockchain is introduced in

M. V. Varada (✉) · W. S. Abdulla · A. Udayakumar · H. Raghu · M. V. Thomas
Vimal Jyothi Engineering College, Chemperi, Kannur, Kerala, India
e-mail: manojkurissinkal@vjec.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_38

this area, people can monitor all the amount with respect to when and where it is being used from the decentralized ledger. Here, we introduce a system over hyperledger fabric technology to track and allocate funds and hence bring down the possibility of corruption.

In the next section, we find the details of the related work and the hyperledger fabric architecture and then, a description of the proposed system with explanation of its working. In the fourth section, a description of our modules is given for this system, and the fifth section also discusses regarding the results. Finally, we talk about the relevance of this application, its future scope and also reach a conclusion.

38.2 Related Work

38.2.1 Application of Blockchain in Government

Governments all over the world are implementing blockchain in different areas. Estonia was one of the first countries to test blockchain in government applications. First by embedding blockchain within data transfer platform X-Road. Second, by testing a "notorization on the blockchain" which was a project combined with BitNation. And thirdly by piloting keyless authentication which is supported by the blockchain distributed ledger technology [2]. In addition, different organizations across the globe as in EU and UK have officially released positive reports on using blockchain [3, 4].

38.2.2 Hyperledger Fabric

Software Development Kit (SDK): SDK allows the users to interact with the hyperledger blockchain network setup and is shown in Fig. 38.1. It creates the channels, and the peer nodes are asked to join the channel. Channels have the ability to partition the transaction visibility for only the participant. Chaincodes are installed in each peer by the SDK.

Membership Service Provider (MSP): An MSP defines its own notion of identity, and the rules by which the identities of the various organizations are governed and authenticated (signature generation and verification) [5].

Peers : Peer nodes can be endorser or committer nodes. It endorses the proposal for transaction or commit blocks of transaction to ledger [6].

Orderer: A block of incoming transactions is created by the orderer and supports the consensus service.

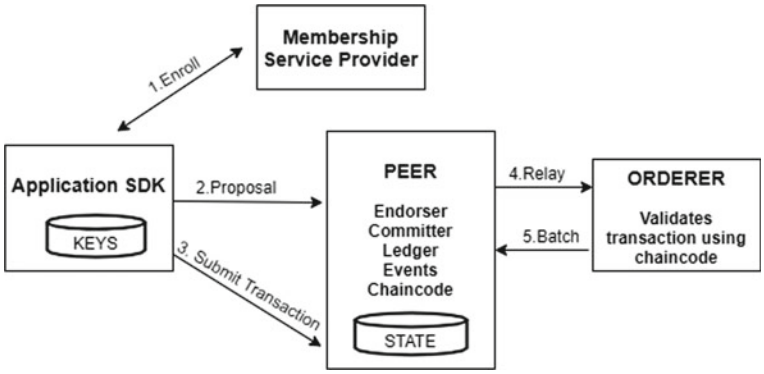


Fig. 38.1 Hyperledger fabric architecture

38.3 Proposed System

In this paper, we propose a system to digitally record the project fund grants. We use the blockchain to record the creation and transfer of digital evidence, and it ensures integrity, credibility and prevents malicious tampering.

38.3.1 Architecture of the System

The participating nodes of the hyperledger fabric are interconnected and can be used for sharing of data and validation, and the transaction behavior of the blockchain and the integrity and trustworthiness of the ledger data can be maintained through the consensus mechanism [7, 8].

1. **Network creation**: A blockchain network is deployed on a local machine or several computers to support the fund transfer process.
2. **Setting parameters**: The status and other parameters required for distributing fund such as status variable, time of transaction, validation, unique applicant ID are set.
3. **Transaction**: Validated applications send transactions; blockchain validators [9], among other things, verify the correctness and the fact that no attempt was made to divert the funds to someone who does not deserve, due to linkability. All transactions that attempt to apply twice are automatically invalidated and are not added to blockchain.
4. **Distribution**: At the end of a particular selection, the project funds are granted to a company or individual. Figure 38.2 depicts the architecture of the proposed system.

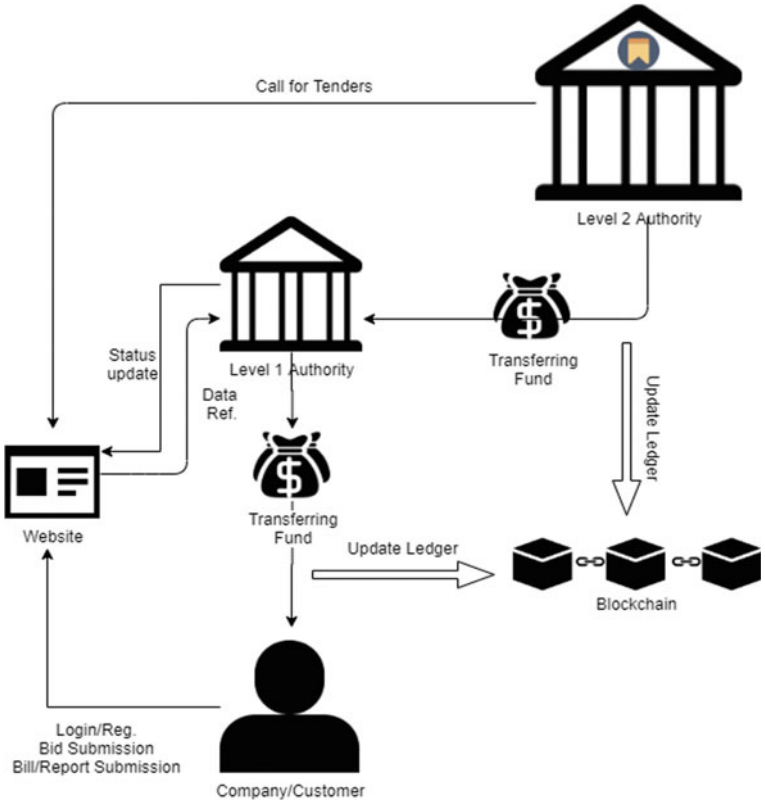


Fig. 38.2 Architecture of the proposed system

38.3.2 Specialized Network

In order to deploy the network, validators and auditors must be identified. Validators are the nodes that form blocks of transactions from the pool of unconfirmed transactions [10, 11]. Auditors are the nodes that store a complete copy of blockchain and do not participate in block mining. Our system uses the hyperledger fabric as it supports the use of channels for transactions and can expand to a large network as compared to Ethereum or other public/private ledger blockchain technologies.

38.3.3 Storage

After defining the node, the validated users are installed. The list of block candidates is formed in the very first block (genesis block) of the network [12]. After defining the users, the transactions are moved to the pool, that is, the funds are granted

to the individuals. After validation based on chaincode, the requested changes in transactions are displayed in blockchain. The default database used by hyperledger fabric to store the states is leveldb. But in order to satisfy the rich queries and the JSON format, Couchdb is used [13].

38.3.4 Requirements

Software requirements	Hardware requirements
Hyperledger fabric	Operating systems: Ubuntu Linux 14.04
Docker	16.04 LTS (both 64-bit)
Git	or Mac OS 10.12
npm	
Node JS	RAM: 8 GB
cURL	

38.3.5 Working of the System

The network is created using hyperledger fabric. A blockchain network is deployed on a local machine or several computers. Figure 38.3 depicts the workflow diagram for fund tracking and management:

1. The company submits bid for the project that is within the budget specified in the call for tender notice and places request for the fund.
2. One of the project proposals is selected by Level 1 authority.
3. Level 2 authority on receiving fund request from Level 1 authority checks for the availability of funds.
4. Level 2 authority transfers the fund to Level 1 authority, and the transaction is recorded in the ledger of hyperledger fabric.
5. Initial payment is released by Level 1 authority.
6. The company submits stage-wise completion report to Level 1 authority.
7. Level 1 authority transfers fund at each stage after verification.

38.4 Modular Description

Call for Tenders: A notification with link is provided in government official portal. The company/customer gets the notification, and they can register using the link. This link will be directed to a Web page.

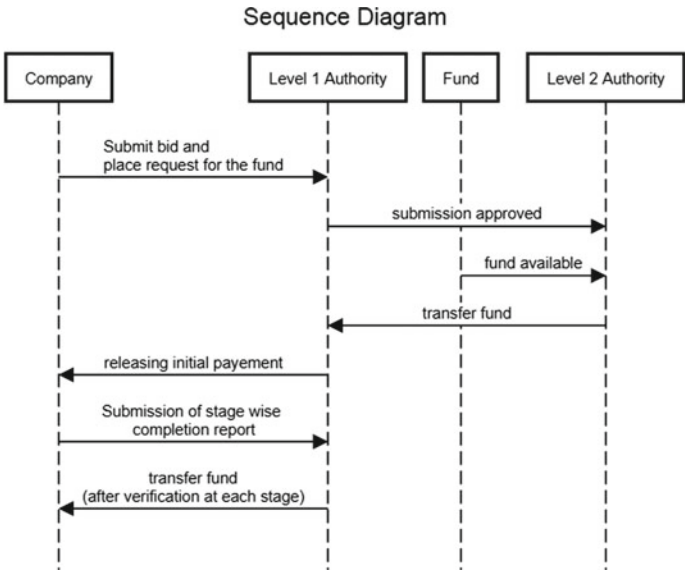


Fig. 38.3 Workflow of the proposed model

Company/Customer Registration: The details needed for company/customer registration are asked in the Web site. The Web site mainly asks for the mail ID, pan number, company ID, project proposal details,mobile number,password and documents.

Verification: Level 1 authority verifies the company/customer by checking their provided ID and profile. In the profile, the required documents must be provided by the company. So, using the documents, Level 1 authority verifies the project proposal, and then if the proposal is valid, then they proceed with further actions. Figure 38.4 shows various organizations involved.

Fund Sanctioning and distribution: The Level 2 authority sanctions the fund. The Level 1 authority, after verifying and checking the requirements of the company or the customer, decides the total amount required for all the project and giving all the information to the Level 2 authority. After looking at the information provided by the Level 1 authority, the Level 2 authority sanctions the fund. The Level 1 authority distributes the fund to company/ customer. The fund distribution is done using the requirements provided by the company or customer.The company/customer gets the fund through the bank account(provided in the profile). As the total fund generated can be viewed for all the participants, corruption is not possible. If there is any complaint regarding the fund distribution, then complaint registration is also possible. The fund safely reaches to the company/customer. This system thus provides trustful environment. The status is displayed on webpage. The transactions would be shown on ledger.

```
! crypto-config.yaml > [ ] OrdererOrgs
1   OrdererOrgs:
2     - Name: Orderer
3       Domain: fundmanagement.com
4       EnableNodeOUs: true
5       Specs:
6         - Hostname: orderer
7   PeerOrgs:
8     - Name: Applicant
9       Domain: applicant.fundmanagement.com
10      EnableNodeOUs: true
11      Template:
12        Count: 2
13      Users:
14        Count: 1
15    - Name: Level1
16      Domain: level1.fundmanagement.com
17      EnableNodeOUs: true
18      Template:
19        Count: 2
20      Users:
21        Count: 1
22    - Name: Level2
23      Domain: level2.fundmanagement.com
24      EnableNodeOUs: true
25      Template:
26        Count: 2
27      Users:
28        Count: 1
```

Fig. 38.4 Screenshot of various organizations involved in the system

38.5 Results and Discussion

Caliper is a framework used to measure blockchain performance, which allows users to test various blockchain solutions with a set of use cases, and get the performance test results under the hyperledger fabric project [14]. The benchmarking criteria in Caliper so far include success rate, transaction/read latency (minimum, maximum, average, percentile) and transaction/read throughput. Figure 38.5 shows the performance matrix of certificates in the hyperledger fabric using this tool.

Name	Succ	Fail	Send Rate (TPS)	Max Latency (s)	Min Latency (s)	Avg Latency (s)	Throughput (TPS)
Certificate Generation	1000	0	192.0	7.00	2.93	4.79	112.7

Fig. 38.5 Performance metrics of certificate

TYPE	NAME	Memory (max)	Memory (avg)	CPU% (max)	CPU% (avg)	Traffic In	Traffic Out	Disc Read	Disc Write
Process	node local-client.js(avg)	42.1MB	43.5MB	10.45	3.30	-	-	-	-
Docker	dev- peer0.org1.fund.com-projectfund-v1	46.1MB	44.5MB	25.01	9.02	1.23MB	456.2KB	0B	0B
Docker	dev- peer0.org2.fund.com-projectfund-v1	44.1MB	44.6MB	0.21	0.00	191B	191B	0B	0B
Docker	peer0.org1.fund.com	289.2MB	290.4MB	60.1	38.2	4.1MB	12.7MB	0B	4.7MB
Docker	ca.org1.fund.com	20.2MB	20.2MB	0.00	0.00	0B	0B	0B	0B
Docker	orderer.fund.com	34.7MB	29.4MB	72.8	21.2	2.4MB	2.5MB	0B	2.9MB

Fig. 38.6 Resource consumption

Resource consumption indicates the resource (CPU, memory, network IO) that is consumed by each Docker container. Each container represents a peer in the blockchain. Figure 38.6 gives an overview of the details regarding the resource consumption of this work.

38.6 Conclusion and Future Work

This model implemented using blockchain helps to provide transparency in all the transactions made by the government. The decentralized ledger makes it possible for verifying the transactions and impossible to alter the transaction information. This means that the fund that is sanctioned can be monitored. Such a blockchain network will surely reduce the corruption that occurs in our society and promote the economic development of the country. It also creates greater trust in government. The key is to land in blockchain technology and build a new system of corruption-free and transparent system.

There is a lot of scope for further document verification to be done digitally using zero validation techniques so that impersonation cannot be done. Any online transaction can be made to verify digitally. Here, the blockchain ensures the security of data. This model can be run on the cloud on the hyperledger fabric where the blockchain can be stored. Also, this project fund system can be expanded to the whole government fund releasing systems such as scholarships, pension funds, relief funds and so on.

References

1. Ismail, L., Hameed, H., AlShamsi, M., Alhammadi, M.S., Al Dhanhani, N.: ICBCT 2019: Proceedings of the 2019 International Conference on Blockchain Technology, Towards a

Blockchain Deployment at UAE University: Performance Evaluation and Blockchain Taxonomy, March 2019, pp. 30–38
2. Chen, S., Shi, R., Ren, Z., et al.: A blockchain based supply chain quality management framework. In: International Conference on E-Business Engineering **2017**, 172–176 (2017)
3. Glaser, F.: Pervasive decentralisation of digital infrastructures: a framework for blockchain enabled system and use case analysis (January 2017). In: Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS-50), Hawaii, January 4–7, 2017. Available at SSRN: https://ssrn.com/abstract=3052165
4. Jun, M.: Blockchain government—a next form of infrastructure for the twenty-first century. J. Open Innov.: Technol. Market Complexity **4** (2018)
5. Wai, K., Htoon, E.C., Thein, N.M.: Storage structure of student record based on hyperledger fabric blockchain. In: 2019 International Conference on Advanced Information Technologies (ICAIT), 6–7 Nov. 2019, pp. 108–113. IEEE, New York
6. Thwin, T.T., Vasupongayya, S.: Blockchain based secret data sharing model for personal health record system. In: 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), 14–17 Aug 2018, pp. 196–201. IEEE, Krabi, Thailand
7. Welcome to Hyperledger Fabric. https://hyperledger-fabric.readthedocs.io/. Accessed 2018
8. Jha, A., Bhattacharjee, R.K., Nandi, M., Barbhuiya, F.A.: BSCI '19: Proceedings of the 2019 ACM International Symposium on Blockchain and Secure Critical Infrastructure: A Framework for Maintaining Citizenship Record on Blockchain, pp. 29–38 (2019). https://doi.org/10.1145/3327960.3332389
9. Ranjan, S., Negi, A., Jain, H., Pal, B., Agrawal, H.: Network system design using hyperledger fabric: permissioned blockchain framework. In: Twelfth International Conference on Contemporary Computing (IC3), 2019, 8–10 Aug. 2019. IEEE, New York
10. Mohite, A., Acharya, A.: Blockchain for government fund tracking using hyperledger. In: 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), 21–22 Dec. 2018, pp. 231–234
11. Welcome to Hyperledger Fabric. https://hyperledger-fabric.readthedocs
12. Benhamouda, F., Halevi, S., Halevi, T.: Supporting private data on Hyperledger Fabric with secure multiparty computation. https://shaih.github.io/pubs/bhh18.html (2018)
13. LevelDB vs. CouchDB. [Online]. Available: http://vschart.com/compare/leveldb/vs/couchdb
14. https://medium.com/@nima.afraz/hyperledger-caliper-explained-and-installation-guide-ubuntu-c38dc16d3dcf

Chapter 39
Random Valued Impulse Noise Detection Using Fuzzy c-Means Clustering Technique



Aritra Bandyopadhyay, Kaustuv Deb, Atanu Das, and Rajib Bag

Abstract Noise detection is critical facet in eliminating impulse noise from digital images. Faulty detection can have deeper consequences in the ultimate outcome of the image filtering procedure. Established detection methods while distinguishing corrupted and non-corrupted pixels, have a tendency to upsurge errors while judging the edge points. In this paper, a detection technique, based on fuzzy c-means clustering algorithm is proposed to deal with above-mentioned issues. Fuzzy c-means clustering is used here to group similar non-corrupted pixels intensities by partitioning them into separate fuzzy clusters. Here, a pixel intensity value belongs to each cluster with certain membership strength. Depending on the maximum membership strength of a pixel element in a certain cluster among all the clusters, the most suitable belongingness of that pixel to that cluster is determined. After this operation, maximum element holding cluster is selected to be the cluster having the detected non-corrupted pixel elements, since there is a plausible tendency of similar non-corrupted pixel elements to group together. Standard measures like "miss" and "false-hit" have been used to compare the proposed method against the prevailing state-of-art methods to determine the noise detection accuracy of the proposed method. Being evaluated upon the aforesaid standard measures, the proposed detector is found superior to all the others in terms of its detection performance.

39.1 Introduction

Noise detection operation, prior to elimination, is an essential task in the arena of image restoration. Every noise type has a detrimental effect on digital images. Impulse noise is one of the types of noise that can arise due to faulty camera operations, atmospheric turbulences, flawed memory locations, or during any image

A. Bandyopadhyay (✉) · K. Deb · R. Bag
Supreme Knowledge Foundation Group of Institutions, Chandannagar, West Bengal, India

A. Das
Netaji Subhash Engineering College, Kolkata, West Bengal, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_39

processing trials. It impeaches negative effect on the image quality and the imminent image processing operations. To resolve this consequence, image restoration is a necessary operation. Impulse noise is categorized into two types: salt and pepper noise (SAPN) [1] and random valued impulse noise (RVIN) [2]. This paper is concentrated on detecting RVIN, which creates random dots in the images having unpredictable pixel intensities in the grayscale range of '0' to '255'. Predicting random changes in numerical intensity values makes RVIN detection a thought-provoking job.

Round the ages, several detectors, some also with removal filters have been established. SD-ROM [3] is one of the methods established in the past, that was having four thresholds to first identify a pixel as noisy or not. After a few years, PSM [4] and TSM [5] filters were proposed. The PSM had detection and removal mechanisms which were gradually applied over frequent iterations that created blurry outputs. TSM, on the other hand, had united, standard median filter and center-weighted median filter to form a detector mechanism with a specific threshold which restricted the performance of the work. Then ACWM [6] was designed to have an adaptive operator to deal with the impulse detection which was again failed to preserve the details of the images over 30% noise density. In the same year, MSM [7] was developed which also used a threshold-based detection technique. In a few years later, another method PWMAD [8] was established. PWMAD demonstrated a vigorous estimator MAD (median of the absolute deviations from the median) to identify corrupted pixels. Both PWMAD and MSM were failed to have adequate result at a noise density upper than 30%. In the meantime, ACWM filter was revised to form ACWM-EPR [9] which used adaptive center-weighted median filter to detect corrupted pixels. This method had some improvement than the previous filters but was unable to preserve image details at a noise density greater than 50%. Afterward, another impulse detector, combined with median filter [10] was introduced. The detector was able to improvise detection by maintaining the image details, but that also failed in high noise density of and over 40%. In the same period, ROAD filter [11] was introduced which had used a completely different technique to identify noise. Local statistics of image was used in it to detect corrupted pixels by identifying the statistical intensity difference from a designated pixel to their utmost alike neighbors. This was followed by another filter ROLD [12] which extended the concept of ROAD by amplifying the intensity differences among corrupted and non-corrupted pixels to make detection stronger. But these two filters altogether dealt up to 50% noise density. In the equivalent year, another filter DWM [13] came with diverse concept to propose an impulse noise detector which was founded on the alterations among the present pixel and its neighbors associated with four key directions which resulted pretty virtuous in respective to detection even at a high noise density of 50%. In recent times a filter [14] detected noisy pixels by means of a sparse depiction of the pixels in a taken window out of the whole image. Another filter [15] used the concept of sparse depiction and combined it with thresholding to form a new detector. Both the detection methods were failed to provide effective outcome over 50% noise densities.

In this paper, fuzzy c-means clustering algorithm is used to accumulate analogous non-corrupted pixels intensities by segregating them into distinct fuzzy clusters and then reliant on most membership strength of a pixel intensity, the residence of the

designated pixel on a particular cluster is determined. Then, most element packed cluster is identified as the best cluster, considered to be having the non-corrupted pixels. Thus, rest of the elements that do not belong to the best cluster are detected as corrupted ones.

The rest of the paper is planned as follows. Section 39.2 demonstrates the proposed methodology. Section 39.3 portrays the results and discussion and the end Sect. 39.4 depicts the conclusion.

39.2 Proposed Methodology

A 512×512 sized noise corrupted image, denoted as IMNC, is considered to perform noise detection process. The image matrix of IMNC is denoted as IMMNC. A flag image matrix FG of same size as IMMNC is created and every pixel's value is replaced by '0'. IMMNC is subdivided into some number of 5×5 windows and the proposed fuzzy c-means clustering implemented noise detection process is applied on each of these windows. The underlying idea is that the non-corrupted pixels present in a window are most likely to form a cluster together as their pixel intensity values are not affected by random noise. In compare to that, all the noise corrupted pixels having noise affected random pixel intensity values are least likely to form a cluster together and instead, they will be scattered in different clusters. Fuzzy c-means clustering is a fuzzy logic-based strong clustering algorithm, has the capability of properly figuring out the similarity pattern existing in the pixel intensity values and thereby forming clusters by grouping similar pixel intensity values together (Fig. 39.1).

Fuzzy c-means clustering assigns a membership strength value, to each pixel intensity value, with which a pixel intensity value becomes a part of a cluster. A membership strength value for each cluster is assigned to a pixel intensity value, and finally, the pixel intensity value is most suitable to belong to the cluster for which the pixel intensity value has the highest membership strength value.

Thus, using fuzzy c-means clustering, it is most likely to find the non-corrupted pixels by tracing out the cluster having the most number of pixel intensity values in it. With this, it also can be said that the pixels, scattered in other clusters having a smaller number of pixel intensity values in them, are detected as noise corrupted pixels. With the purpose of forming four clusters, fuzzy c-means clustering is applied on each 5×5 window. CL_i, $i = 1$–4, denotes a cluster of pixel intensity values present in a window and CLM denotes the CL_i, $i = 1$–4, which has the most number of pixel intensity values in it. PV_N denotes the pixel intensity value of the center pixel, denoted as N, of the window. N's position in IMMNC is (r, c). The proposed detector detects N as a noise corrupted pixel if $\mathrm{PV}_N \notin \mathrm{CLM}$; otherwise N is considered as a non-corrupted pixel. If N is found corrupted, then the flag matrix's position $\mathrm{FG}[r][c]$ corresponding to N is replaced by '1' and otherwise is replaced by '0'.

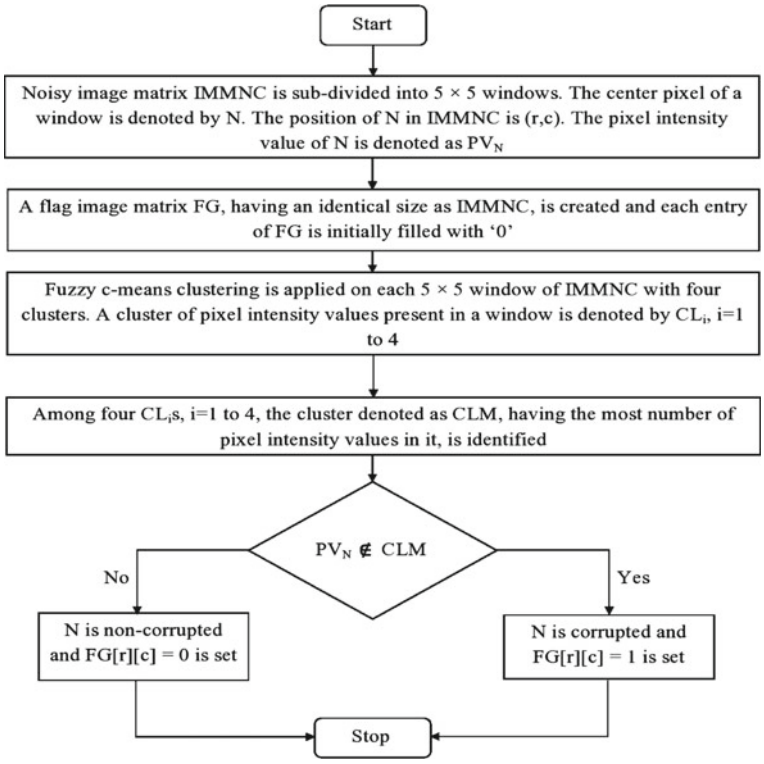


Fig. 39.1 Flowchart of proposed detection technique

39.3 Results and Discussion

The proposed work is carried out in MATLAB 2018a software, mounted on Windows PC having Core i5 processor and 8 GB of RAM. The proposed method's noise detection ability is compared in counter to thirteen state-of-art methods. The proposed method is applied on seven common test images: Lena, Goldhill, Baboon, Barbara, Bridge, Boat and Fingerprint each having a size of 512 × 512, to judge the outcome of its detection performance with respect to image assortment.

Noise detection accuracy of the proposed method is primarily shown in Fig. 39.2. Three 5 × 5 sample windows from the image matrix IMMNC for three different images are depicted here to highlight the outcome of the work. These three windows are the original image sample window, the corresponding same positional noisy image and the flag image window, respectively. Noisy image window is the RVIN affected image matrix's sample window which displays the noise exaggerated section. Flag window shows the refection of the activity of the detection procedure through the demarcation of the corrupted pixels with flag value '1' and non-corrupted pixels with flag value '0'. Rectangle outline boxes are taken to mark the wrongly

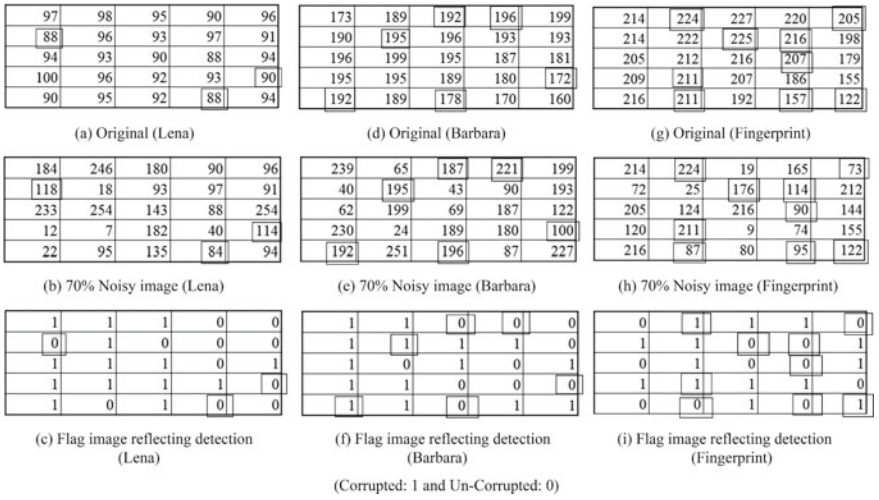


Fig. 39.2 Proposed detection performance measure

detected pixels. Lena, Barbara and Fingerprint are used here to demonstrate the detection result at 70% noise density. From Lena image perspective, due to less variability of its pixel intensities, the detection outcomes are excellent, where only three marked errors found examining the sample windows. Barbara image is more complex image than Lena due to its image depth and because of this, the detection result for Barbara image is respectable but not as good as Lena. Texture image, Fingerprint is more intricate than both Lena and Barbara in terms of image profundity. Due to the high erraticism in the pixel's intensities, it is very tough to carry out detection procedure of Fingerprint image. Still, our detection gave reputable result for Fingerprint image. Moreover, reputable results shown in Fig. 39.2 are obtained at sufficiently high 70% noise density, which portrays the fact that the quality of the detection algorithm is good.

Table 39.1 validates the detection meticulousness of the proposed algorithm by showing the number of erroneous pixels in comparison with other state-of-art filters by means of "miss" and "false-hit". The term "miss" signifies the number of actually corrupted pixels which are detected as non-corrupted by the proposed detector and "false-hit" implies the number of actually non-corrupted pixels which are detected as corrupted by the proposed detector. From Table 39.1, it is identifiable that SD-ROM and ACWM have a very low false-hit number which is exceptional but both of them have much more higher miss value. The miss value of the proposed method is less but greater than some methods like DWM, ROLD-EPR, SAIKRISHNA, and DEKA in some cases. But if we judge in terms of the total noise (miss + false-hit), then the proposed filter wins the battle against all methods at 50% and above noise densities. So, the proposed detector exhibited admirable performance at high noise densities.

Table 39.2 portrays the "miss" and "false-hit" values of the proposed detector for seven different test images at diverse noise densities. Table 39.2 shows the detection

Table 39.1 Comparison of various detection methods for Lena image aimed at RVIN

Methods	40%		50%		60%	
	Miss	False-hit	Miss	False-hit	Miss	False-hit
SD-ROM [3]	22,842	411	32,566	998	45,365	2651
PSM [4]	23,461	4773	29,396	8550	35,617	15,152
TSM [5]	18,703	3974	12,035	26,806	15,492	33,029
ACWM [6]	16,052	1759	23,683	2895	32,712	7644
MSM [7]	16,582	7258	20,857	10,288	26,169	15,778
PWMAD [8]	11,817	9928	14,490	15,003	17,760	19,577
ACWM-EPR [9]	13,657	6192	13,868	12,693	23,793	17,573
LUO [10]	16,706	3678	23,707	6410	38,512	12,724
ROAD-EPR [11]	13,476	8079	13,771	10,055	17,212	9330
ROLD-EPR [12]	12,010	7404	13,329	7761	14,967	9109
DWM [13]	9512	7761	9514	11,373	12,676	12,351
SAIKRISHNA [14]	31,755	1215	39,041	2047	45,908	4260
DEKA [15]	21,061	2063	24,903	3195	32,722	5047
PROPOSED	12,673	5645	13,841	5965	15,458	6355

Table 39.2 Detection results of the proposed method for various images at diverse noise densities

Images	40%		50%		60%	
	Miss	False-hit	Miss	False-hit	Miss	False-hit
Lena	12,673	5645	13,841	5965	15,458	6355
Goldhill	13,789	6958	14,598	7215	15,948	7853
Baboon	15,789	7856	15,998	8165	16,523	8639
Barbara	18,211	9562	19,365	10,245	21,548	11,149
Bridge	15,964	8102	16,235	8614	17,023	9017
Boat	16,245	8455	16,892	8824	17,226	9324
Fingerprint	23,548	11,269	25,128	13,019	27,693	15,238

accuracy of the proposed work for diverse images having dissimilar image depth and features. Lena, Goldhill and Baboon are the less intricate images where the proposed filter gives lower miss and false-hit results. In the case of Bridge, Boat and Barbara images, the proposed work also demonstrated good but higher total errors. Only Fingerprint image is the exception as it is having texture information that creates higher complications. It shows formidable false-hits but higher miss values. But altogether, if we judge the total noise, then all the images, except the Fingerprint, establish excellent outcomes.

As the proposed method uses fuzzy c-means clustering for detection purpose, a predetermined cluster size '4' is used. The cluster size is chosen on trial and error basis. Generally, as per the need of the algorithm, to properly find the cluster having

maximum number of elements in it, minimum of three clusters are needed to be formed. Then only the homogenous non-corrupted pixels in a sample 5 × 5 window can form a maximum number of elements containing cluster. Moreover, too many clusters, like taking '6' clusters, are also not pertinent, as it can hamper the motive of the use of fuzzy c-means clustering. In Table 39.3, we have shown the "miss" and "false-hit" values taking cluster size as 3, 4, and 5 for seven test images at 60% noise density. It can be seen that the total erroneous numbers for cluster size 3 and 5 are a bit more. It justifies the reason to take cluster size as 4 for our detector.

Table 39.4 demonstrates the variation of window sizes that can be considered. The window size begins with 3 × 3 to keep it odd. Here 5 × 5 window size is taken for the proposed detector as compared to other mentioned window sizes, on the basis of trial and error to get the least error values. The table shows results of "miss" and "false-hit" values for 3 × 3, 5 × 5, 7 × 7 and 9 × 9 window sizes for three images at different noise densities. It can be seen that for 3 × 3 and 7 × 7 window sizes, the error values are deviated by reasonable but small amount. But those windows produced greater miss and false-hit values than the implemented 5 × 5 window. The

Table 39.3 Detection results by varying the cluster sizes for diverse images at 60% noise density

Images	Cluster size					
	3		4		5	
	Miss	False-hit	Miss	False-hit	Miss	False-hit
Lena	15,896	6539	15,458	6355	16,126	6896
Goldhill	16,114	7921	15,948	7853	16,549	8194
Baboon	16,658	8859	16,523	8639	17,011	9254
Barbara	22,021	12,169	21,548	11,149	22,716	12,532
Bridge	17,263	9368	17,023	9017	17,796	9736
Boat	17,985	9869	17,226	9324	18,161	10,254
Fingerprint	28,691	16,147	27,693	15,238	28,998	16,735

Table 39.4 Results of the proposed detector by varying the window size for diverse images

Images	Window size	50% ND		60% ND	
		Miss	False-hit	Miss	False-hit
Lena	3 × 3	14,019	6112	15,797	6658
	5 × 5	13,841	5965	15,458	6355
	7 × 7	13,998	6012	15,536	6521
	9 × 9	14,652	6524	15,997	6754
Barbara	3 × 3	19,721	10,469	21,765	11,324
	5 × 5	19,365	10,245	21,548	11,149
	7 × 7	19,659	10,411	21,694	11,206
	9 × 9	19,861	10,524	21,814	11,598

least window size of 3×3 has failed more at high noise densities, because, maximum number of pixels might be noisy and thus small numbers of non-corrupted pixels are left to form cluster. The 7×7 and 9×9 windows are also been tested. Those window sizes produced inferior results than 5×5. This validates the choice of 5×5 window and the superior results are reflected by the designated Table 39.4. Altogether the entire results portrayed superior performance of the proposed detector, even at high noise densities.

39.4 Conclusion

A random valued impulse noise detection technique based on fuzzy c-means clustering is proposed in this paper. The results justified the fact that use of fuzzy c-means clustering technique can be very useful to group analogous non-corrupted pixels to segregate them from the corrupted ones. Metrics like "miss" and "false-hit" validated the superior performance of the proposed detector at high noise densities. In future, another clustering method called K-medoid can be explored for detection operation.

References

1. Banerjee, S., Bandyopadhyay, A., Bag, R., Das, A.: Sequentially combined mean-median filter for high density salt and pepper noise removal. In: International Conference on Research in Computational Intelligence and Communication Networks, pp. 21–26. IEEE (2015)
2. Banerjee, S., Bandyopadhyay, A., Mukherjee, A., Das, A., Bag, R.: Random valued impulse noise removal using region based detection approach. Eng. Technol. Appl. Sci. Res. **7**(6), 2288–2292 (2017)
3. Abreu, E., Lightstone, M., Mitra, S.K., Arakawa, K.: A new efficient approach for the removal of impulse noise from highly corrupted images. IEEE Trans. Image Process. **5**(6), 1012–1025 (1996)
4. Wang, Z., Zhang, D.: Progressive switching median filter for the removal of impulse noise from highly corrupted images. IEEE Trans. Circuits Syst. II Analog Digit. Signal Process. **46**(1), 78–80 (1999)
5. Chen, T., Ma, K.K., Chen, L.H.: Tri-state median filter for image denoising. IEEE Trans. Image Process. **8**(12), 1834–1838 (1999)
6. Chen, T., Wu, H.R.: Adaptive impulse detection using center-weighted median filters. IEEE Signal Process. Lett. **8**(1), 1–3 (2001)
7. Chen, T., Wu, H.R.: Space variant median filters for the restoration of impulse noise corrupted images. IEEE Trans. Circuits Syst. II Analog Digit. Signal Process. **48**(8), 784–789 (2001)
8. Crnojevic, V., Senk, V., Trpovski, Z.: Advanced impulse detection based on pixel-wise MAD. IEEE Signal Process. Lett. **11**(7), 589–592 (2004)
9. Chan, R.H., Hu, C., Nikolova, M.: An iterative procedure for removing random-valued impulse noise. IEEE Signal Process. Lett. **11**(12), 921–924 (2004)
10. Luo, W.: A new efficient impulse detection algorithm for the removal of impulse noise. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **88**(10), 2579–2586 (2005)
11. Garnett, R., Huegerich, T., Chui, C., He, W.: A universal noise removal algorithm with an impulse detector. IEEE Trans. Image Process. **14**(11), 1747–1754 (2005)

12. Dong, Y., Chan, R.H., Xu, S.: A detection statistic for random-valued impulse noise. IEEE Trans. Image Process. **16**(4), 1112–1120 (2007)
13. Dong, Y., Xu, S.: A new directional weighted median filter for removal of random-valued impulse noise. IEEE Signal Process. Lett. **14**(3), 193–196 (2007)
14. Saikrishna, P., Bora, P.K.: Detection and removal of random-valued impulse noise from images using sparse representations. In: International Conference on Image Processing, pp. 1197–1201. IEEE (2013)
15. Deka, B., Handique, M., Datta, S.: Sparse regularization method for the detection and removal of random-valued impulse noise. Multimed. Tools Appl. **76**(5), 6355–6388 (2017)

Chapter 40
Distress-Level Detection Using Deep Learning and Transfer Learning Methods



Srijha Kalyan, Hamsini Ravishankar, and C. Arunkumar

Abstract The mental health of young adults and teenagers proves to be vital for having a flourishing life. Neglecting the issue of mental health can lead to anxiety, stress, and depression. These problems need to be addressed in the early stages to ensure better mental health for young adults. This paper provides a comprehensive study by leveraging natural language processing and deep learning techniques to detect depression by examining the relationship between language usage and the psychological characteristics of the communicator. A private dataset DAIC-WOZ has been used to predict the levels of depression. It includes audio, video, and textual information from 189 interviews. This paper focuses on the textual analysis of detecting depression from transcripts of the interviews conducted by an animated virtual interviewer. Due to the disadvantages faced by the context-independent nature of GloVe embeddings, the proposed approach uses transfer learning techniques such as ELMo, ULMFit, and BERT for predicting depression severity from transcripts. Furthermore, a Python Web application has been deployed to identify negative sentiments and depression severity from sentences inputted by the user. The proposed approach uses an ensemble learning method for the application to provide better predictions for classifying the texts into levels of depression. Hence, the inferences made in this paper can be extrapolated to other all demographics around the world to help detect depression in textual data as the algorithms and techniques used are all-encompassing..

40.1 Introduction

With the rise in easy access to social media and the Internet, the psychological health of young adults and teenagers is prone to getting affected. Neglecting to address this issue can lead to depression and in severe situations, even suicidal thoughts.

S. Kalyan · H. Ravishankar · C. Arunkumar (✉)
Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: c_arunkumar@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_40

Early detection of a quick and successful automated diagnosis of depression could be of a great help. Depression causes cognitive and motor changes that affect speech production. Reduction in verbal activity productivity, prosodic speech irregularities, and monotonous speech have all been shown to be symptoms of depression.

Traditional machine learning techniques have not been able to capture the context of the words and the sentiments of the textual information to detect distress. Deep learning techniques have been incorporated to detect depression from textual data, speech and video data. Recent developments have mainly been focused on the detection of depression through multi-modal fusion by combining various modalities or types of information and deep learning methods using the audio, video, and textual information provided in the database [6, 7].

This paper focuses only on transcripts of the clinical interviews to detect the depression severity in the patients. In summary, the main contributions can be described as follows. Firstly, deep learning models using GloVe word embeddings along with LSTMs, GRUs, TextCNNs, and BiLSTMs have been implemented. To draw attention to the important words used by depressed patients, BiLSTM with Attention Layer has been proposed. Secondly, due to the disadvantages of the context-independent nature of GloVe word embeddings, Embeddings of Language Model (ELMo), a context-dependent model has been used. Thirdly, in order to understand the significance of transfer learning in NLP, Universal Language Model Fine Tuning (ULMFit) and Bidirectional Encoder Representation from Transformers (BERT) have been proposed for predicting the depression severity. Finally, a Python Flask Application using has been incorporated by leveraging ensemble method of averaging the predictions from all the algorithms that use GloVe word embeddings.

40.2 Related Methods

The following section provides a review of the literature related to the development of various deep learning models used to detect whether a person is suffering from distress or not. Ringeval et al. [1] in the "Real-life Depression, Recognition Workshop and Challenge" presented an approach where the baseline severity of depression was measured using random forests. The proposed approach showed that by combining the regression outputs of the unimodal random forest regressors, the fusion of audio and video modalities can be achieved. Dinkel et al. [2] proposed an approach that used multi-task modeling with pre-trained embedding of sentences, namely Elmo and BERT, to classify depression based on text. The Elmo and BERT models showed a connection between short sounds like "um" and the output of the script, suggesting that in order recognize depression, the text's behavioral aspects must be focused rather than material.

Further developments were done by Li et al. [3] where they proposed a model that used all data from audio, video, and text to extract the features. The proposed multimodel approach was based on a random forest regressor that predicted PHQ-8 scores ranging from 0 to 25, providing an indication of the extent of the stress/tension.

This work also used Covarep toolbox to get features from the audio. For getting features from video, the Facial Action Coding System was used. Moreover, Linguistic Inquiry and Word Count and AFINN toolbox were used for text data. Qureshi et al. [4] presented an early fusion approach that used neural networks and incorporated acoustics, visuals and textual modalities. It focused on the use of modality encoders, fusion sub-network, tensor fusion layer, Attention fusion sub-network and PHQ-8 Performance regression sub-network to achieve this. The results showed that integrating all modalities helped to provide the correct estimate of the degree of depression. In the regression phase, text modality was important and their proposed CombAtt architecture was much better than many other approaches.

In "Multi-level Attention network using text, audio and video for Depression Prediction" by Ray and Kumar [5] an early fusion-based multi-level attention network was proposed that integrated audio, video and text modalities to predict depression severity. The attention network gave maximum weight to the text modality and almost equal weight to the audio and video modes, was observed. Samareh et al. [6] adopted an input-specific classifier for each modality and a decision-level fusion module for predicting the final outcome. The paper used random forest to convert features into predictive scores for each biomarker modality, while those scores were further combined in a confidence-based fusion process to make the final prediction of PHQ8. Rohanian et al. [7] presented a model that learned depression indicators from all types of data in transcripts, including audio, video and lexical content. The paper used word-level multimodal fusion as a gating mechanism, used feed-forward highway layers. Venkataraman and Parameswaran [8] proposed a comprehensive study of facial expressions related to depression and their extraction methods. Happy faces dataset created by the JAFFE database was used. The facial characteristics for each face were then found using a 40 filter bank with Gabor filter. The functionality concatenated vector for each image was generated to form a training feature set. Bhaskar et al. [9] presented a new classification system for identifying emotions in human speech utterances that have been suggested. This method exploited corresponding audio and textual features. Multiclass SVM was used for classification of the emotions.

40.3 Proposed Approach

40.3.1 Dataset and Feature Selection

The Distress Analysis Interview Corpus (DAIC-WOZ) dataset used in the Audio/Visual Emotion Challenge (AVEC 016) challenge has been used. It includes text, audio, and video information of 189 interviews designed to enable the treatment of emotional disturbance, nervous breakdown, and other mental health issues. The participants were asked to do a written Patient Health Questionnaire (PHQ8) test to identify depression levels. The scores were treated from (0–25) in bins separating them into levels of depression: none (0), mild (0–5), moderate (5–10), moderately severe (10–15), severe (15–25).

Feature extraction has been carried out for the verbal responses of each participant as given in the transcript file. From previous research work, it has been noticed that there are changes in the prosodic tones in a depressed person than a normal person. Also, they speak for a lesser amount of time with longer pauses. Consequently, from the words spoken by the participants which was subsequently made into a transcribed format, the total words uttered by the participant are calculated, and the average number of words has also been extracted in a similar manner. In addition, the time taken to complete the interview has also been taken into account. Furthermore, the count of the filler words such as "uh", "um" etc. used by the participants has been calculated. The number of times the participant had uttered in terms of first person by using a lot of "I" s are counted. An analysis of the frequency of the number of positive and negative words such as hate, horrible, miserable, etc. used by the participants has also been done. Harvard positive and negative word dictionary was used, and the words used in the conversations were extracted the frequency of these words has also been found.

40.3.2 Methodology

The following steps have been used for pre-processing the data and producing tokenized words from the transcript file which has been shown in Fig. 40.1.

1. Participant answers were subjected to stopword elimination and optionally stemming/lemmatization, and translated into a list of words.
2. Word tokenization was done, after all the words used in the responses has been collected.
3. By using a sliding window of size 10, the set of terms thus obtained is split into several sets.
4. The corresponding tokens list was padded with zeros for lists (after windowing) that have less than 10 characters.
5. For our layout, we had used GloVe embeddings, specific embeddings featuring 6 billion tokens, 400k word vocabulary and 100 dimensional vectors. A 2-D array was created which stores the vectors in the response vocabulary for each word (after removing the stopwords).



Fig. 40.1 Data preprocessing

40.3.3 Transfer Learning Models Implemented

Embeddings of Language Model (ELMo) Due to the disadvantages faced by the context-independent nature of GloVe embeddings, ELMo embeddings has been incorporated. ELMo is a deep contextualized representation of words that models both dynamic word-utilization features and the variations in semantic understandings. Developed by Allen NLP, the state-of-the-art pretrained model is context dependent. ELMo uses two bidirectional LSTMs in training, so the language paradigm knows both the succeeding and the preceding words in the expression fed to the model. This consists of a two-layer, bidirectional LSTM backbone.

Universal Language Model Fine Tuning (ULMFiT) ULMFiT works using the concept of inductive transfer learning. Inductive transfer refers to a learning mechanism's ability to improve performance on the present task after learning a different but related concept or skill on a prior task. Language model pretrained on the WikiText-103 dataset has been used to train the target task which is then fine-tuned for the classification model. In the proposed approach, the pre-trained language model is fine-tuned on the target DAIC-WOZ dataset.

Bidirectional Encoder Representations from Transformers (BERT) BERT is an open-sourced natural language processing pre-training model developed 2018. It uses L = 24 hidden layers, a hidden size of H = 1024, and A = 16 attention heads. On Wikipedia and BookCorpus, this model has been pre-trained for English. Inputs have been "uncased," i.e., the text has been lower-cased into word fragments before tokenization, and any accent marks have been removed. For preparation, the random input masking has been applied to word pieces independently.

40.3.4 Ensemble Methodology in Flask Application

A web application has been created using the light-weight Flask Framework. The pre-processed tokenized words are converted to word embeddings using the pre-trained GloVe word embeddings. The word embeddings are trained with each model to produce predictions followed by ensemble averaging method to produce better predictions. The ensemble predictions are then written into the Flask server and the predictions are the outputs based on the user's input sentences.

40.4 Results and Discussion

40.4.1 Model Performance Analysis

After running 30 epochs for the GloVe word embeddings + LSTM model with Adam optimization and categorical cross entropy loss, a test accuracy of 94.7% was obtained

```
All is going right with the party, I'm happy to know new people
none
I want an ice cream and have some fries for lunch
mild
I'm afraid of losing my work, I don't have any money
severe
I'm worried about my future, I'm afraid of it
severe
I am a graduate student
none
I am getting married
severe
This party is great, I know lots of people
none
I miss my parents, brothers and sisters
mild
I detest my horrible job
moderate
suicide
severe
```

Fig. 40.2 Output results for LSTM+GloVe model

without balancing the dataset. After the balancing the dataset, the accuracy increased to 98.6% with a test loss of 4.3%. Figure 40.2 represents the output results for the GloVe+LSTM model obtained for the test data.

Due to the significance of context dependency in sentences, ELMo was incorporated and gave an accuracy of 73.1%. Experimental results showed that the model with balanced data worked better than the imbalanced data. After implementing transfer learning using ULMFiT and BERT, it was observed that both the models performed well by obtaining an accuracy of 93.2% and BERT model with 82%.

Confusion matrix, a visual representation of the performance of all the models, has been incorporated. It depicts the number of predictions done by the model that have been classified correctly or incorrectly as true positives, false positives, true negatives, and false negatives. The precision, recall, and F1-scores have been calculated for each model from the confusion matrix obtained.

Precision and Recall: Precision is the measurement of the positive samples that are actually correct. It can be calculated by implementing the formula in Eq. (40.1).

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{40.1}$$

Recall is the proportion of the actual positives that have been correctly identified. Equation (40.2) represents the formula used to calculate recall.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{40.2}$$

Table 40.1 Classification report of the language models

Models	Accuracy	Precision	Recall	$F1$ score
GloVe + LSTM (imbalanced)	0.94	0.93	0.94	0.95
GloVe + LSTM (balanced)	0.98	0.97	0.98	0.96
GloVe + 2 LSTMs	0.82	0.82	0.85	0.81
GloVe + BiLSTM	0.68	0.67	0.64	0.65
GloVe + Bi-GRU	0.93	0.94	0.89	0.93
GloVe + Text CNN	0.82	0.88	0.86	0.87
CNN + LSTM (Hybrid Model)	0.66	0.70	0.67	0.68
BiLSTM Attention Model	0.86	0.90	0.88	0.87
ELMo	0.74	0.74	0.73	0.75
ULMFiT	0.93	0.93	0.94	0.94
BERT	0.84	0.85	0.84	0.83

For a multi-class classification problem, macro-averaging and micro-averaging are performed for calculating precision and recall. In macro-averaging, the average of all the various classes is calculated, whereas for micro-averaging the precision and recall are calculated separately for each class and later added to produce the final results of precision and recall.

$F1$-Score: $F1$-Score has also been calculated as a performance measure for all the models. It is the harmonic mean of precision and recall. The $F1$-score provides a realistic measure of the test data performance using both precision and recall. The formula for calculating $F1$-Score has been shown in Eq. (40.3) . According to Table 40.1, a classification report has been computed for all the language models that have been implemented.

$$F1 - \text{Score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \tag{40.3}$$

40.5 Conclusion and Future Work

The objective of detecting depression through textual information involved in understanding the verbal and sentimental usage of words by the participants. In this paper, it was achieved by using advanced NLP techniques. From the analysis of the dataset,

it was found out that there was an imbalance in the dataset that could have caused a bias toward the mild level of depression. It was solved by using random undersampling. Experimental results showed that the model with the balanced data worked better than the imbalanced data. The BERT model gave good results, and the GloVe + LSTM model outperformed all other models. Moreover, work needs to be done to explore Explainable AI in order to understand decision making processes in the mild/severe levels which will help in understanding the models better because understanding the output is pivotal for researchers and professionals alike. Since the scope of the project has been restricted to textual information, the scope of the project can be widened to explore on the analysis of audio and video features using deep learning techniques. Furthermore, detecting depression on social media such as Facebook or Twitter can also be done in the future.

Declaration We, authors declare that we have taken required permissions from the respective person/authority to use the image/dataset in our work. We shall be responsible in case any dispute arises in the future with regard to this.

References

1. Ringeval, F., Schuller, B., Valstar, M., Gratch, J., Cowie, R., et al.: AVEC 2017—"Life depression, and affect recognition workshop and challenge". In: 7th International Workshop on Audio/Visual Emotion Challenge, AVEC'17, co-located with the 25th ACM International Conference on Multimedia, MM, Oct. 2017, pp. 3–9, Mountain View, USA
2. Dinkel, H., Wu, M., Yu, K.: Text-based depression detection: what triggers an alert. In: MoE Key Lab of Artificial Intelligence SpeechLab. Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China (2019)
3. Li, B., Zhu, J., Wang, C.: Depression severity prediction by multi-model fusion. In: China The Third International Conference on Informatics and Assistive Technologies for Health-Care. Medical Support and Wellbeing, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China (2018)
4. Qureshi, S.A., Hasanuzzaman, M., Gael Dias, S.S.: The Verbal and Non Verbal Signals of Depression-Combining Acoustics. Text and Visuals for Estimating Depression Level. IEEE, New York (2019)
5. Ray, A., Kumar, S., Reddy, R., Mukherjee, P., Garg, R.: Multi-level Attention network using text, audio and video for depression prediction
6. Samareh, A., Jin, Y., Wang, Z., Chang, X., Huang, S.: Predicting depression severity by multimodal feature engineering and fusion. In: AAAI Conference (2018)
7. Rohanian, M., Hough, J., Purver, M.: Detecting depression with word-level multimodal fusion. In: Interspeech (2019)
8. Venkataraman, D., Parameswaran, N.S.: Extraction of facial features for depression detection among students. Int. J. Pure Appl. Math. **118**, 455–462 (2018)
9. Bhaskar, J., Sruthi, K., Nedungadi, P.: Hybrid approach for emotion classification of audio conversation based on text and speech mining. In: International Conference on Information and Communication Technologies (ICICT 2014)

Chapter 41
Fingerprint Genuinity Classification



Meega Annu Jacob and Nirmal Varghese Babu

Abstract A dataset that consists of both contactless and contact-based fingerprints was collected using various instruments like sensors, images, etc. Sensor Interoperability issue was handled using this system. Implementation of contactless fingerprint technologies widely depends on the progressive capability to match contactless 2D fingerprints with contact-based fingerprint databases. Contactless 2D fingerprint identification is more safe and it allows deformation-free imaging for advanced accuracy. Here, an Auto Stack Encoder for accurately matching contact-based and contactless fingerprint images was proposed. Auto Stack Encoder is a neural networks for which the input is similar as the output as they work by reducing the input into a latent-space representation and then again constructing the output from the corresponding image. This Auto Stack Encoder is trained using the images of fingerprint and features and classified as genuine (1) or fake (0). Therefore, an additional robust thin-plate spline model from the concatenation of deep extracted feature vectors made from different networks. This paper shows the Unsupervised learning approach for the Fingerprint Genuinity Classification of Contactless and Contact based fingerprints using the Auto Stack Encoder.

41.1 Introduction

Fingerprint recognition technology [1] is unique of most methods for human identification, say authentication. Contactless 2D/3D [2, 3] fingerprint identification systems is a newly introduced method to identify the limits of customary systems uses the contact-based fingerprint [4]. Development of the advancement in abilities to match contact-based fingerprint images with contactless fingerprints [5, 6] is critical for the success of technologies as millions of fingerprints in huge databases have been acquired using contact-based technologies.

M. A. Jacob · N. V. Babu (✉)
Department of Computer Science and Engineering, Amal Jyothi College of Engineering, Kanjirappally, Kerala, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_41

Comparing the contact based 2D fingerprint [7, 8] images with contactless fingerprints are acute for the success of contactless 2D fingerprint technologies, which offer more safe and deformation less acquisition of fingerprint features. The minutia feature is supposed to be the precise fingerprint feature which has its competence and dependability in recognizing fingerprints. Among the contactless and contact-based fingerprint sensors, it is problematic to collect minutiae features and to guarantee their images from two such sensors [4].

41.2 Related Works

To identify the deformations in the contact-based fingerprints, some approaches have been introduced. Cappelli et al., states a plastic distortion model to address the non-linear deformations in contact-based fingerprints. In Ross and Jain [9], fingerprint sensor interoperability problem for two different flat (plain) fingerprint sensors and they examined the extent of degradation in the matching accuracy when images from two such different sensors are matched. In [10], the thin-plate spline model was introduced to identify the deformation in fingerprints. Local minutiae are found out for the deformation correction and the global minutiae are used for further refining the correction. Global minutiae [11] features are used to match the fingerprints.

Ross et al. [12] proposed a deformation model to identify the fingerprint impressions deformation. They introduced a thin-plate spline model on multiple fingerprint impressions from the same fingers and also found out the average deformation of the fingerprints. In Ross and Nadgir [13], proposed a thin plate spline calibration model to image the fingerprint sensor interoperability. This average deformation model was learnt to correct the deformations of fingerprints from the different sensors.

41.3 Proposed System

The various modules that include in the proposed method are:

- Data Collection
- Fingerprint Preprocessing
- Feature Extraction
- Classification using Auto Encoder
- Results Evaluation.

The block diagram of the proposed approach using Auto Stack Encoder is illustrated in Fig. 41.1. Model is trained with the contact-based fingerprint images with several contactless fingerprints from the same fingers. For training the contact-based fingerprint impression deformation correction model, RTPS Model was used. The unknown contact-based fingerprints is into deformation correction using the Deformation correction model for the better matching with the contactless fingerprint. This

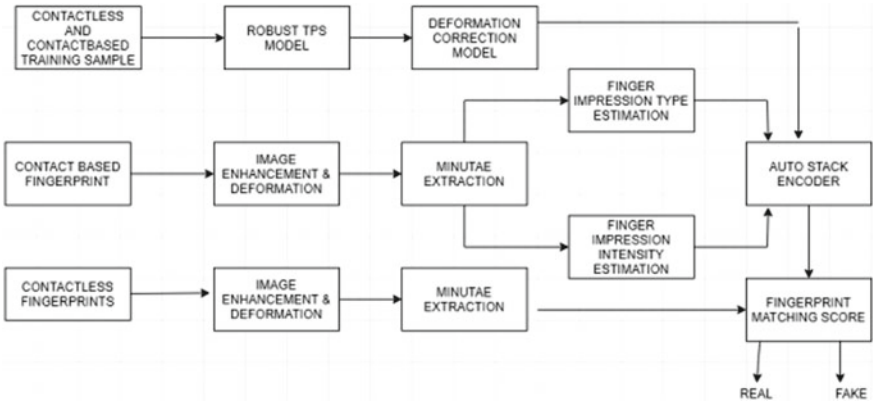


Fig. 41.1 Proposed system

approach estimates the contact-based fingerprint impression type by comparing its image centre. Fingerprint impression intensity is computed from the intensity of grey levels in the fingerprint image. Based on corresponding impression type and intensity, the deformation of each of the contact-based fingerprint is corrected with the help of Auto Stack Encoder. So, the paper is using Auto Stack Encoder to create deep learning recognition system for Contact based and Contactless fingerprint. After the classification of the fingerprints (fake or genuine), the performance of the particular system was evaluated and the accuracy of the system was calculated using the confusion matrix created. Figure 41.1 shows the proposed system architecture.

PolyU dataset was used. It consists of 1800 2D contact-based fingerprint images and 1800 2D contactless fingerprints from 300 different fingers. The contactless 2D fingerprint images are captured using a low-cost camera and lens. 2D contact-based fingerprints were collected using URU 4000 fingerprint reader with FBI/NIST standards. In this work, the dataset contains 150 samples, for each finger, there are 2 contactless fingerprint samples and 4 corresponding contact-based fingerprint samples. Only 100 images of the dataset were used for the training and 50 for testing. The resolution of 2D contact-based fingerprint images and contactless fingerprint was respectively 1400*900 and 356*328 with 500 dpi, respectively.

Fingerprint Preprocessing Because of the essentiality of contact-based and contactless [14] sensing technologies, the images from the corresponding finger using such sensors looks different. Therefore, it is necessary to introduce fingerprint image preprocessing. The same scale ratio or size was ensured by the down sampling operation The region of interest [2] in the images is automatically removed with the centre position of fingerprint impression. For increasing the similarity between the two cross sensor fingerprints from the same subject/finger, Histogram equalization is applied on all contactless and contact-based fingerprints.

Feature Extraction The minutiae features [15] are most correct and they are employed in almost all the fingerprint technologies available nowadays. The drawback of the degradation in matching accuracy [16] is due to minutiae missing is more serious in the contactless fingerprints because the nature of contactless allows a high degree of freedom in the collection under the limited field of view and distortions in regions away from the image center. Therefore, it is unanimous to introduce the minutiae feature [10, 13] and respective ridge feature to improve the learning for the joint fingerprint feature from the two sensors. For each fingerprint, minutiae map is generated by marking the minutiae location [12], direction and quality on the image. Each minutiae point is marked using a solid circle of radius equal to 2 pixels, to avoid the overlapping.

Classification: Auto Stack Encoder The main motive is to match the fingerprints collected from different sensors accurately. Auto Stack Encoder [17, 18] classifies the contactless and contact-based fingerprints [17–23] and it finds the accuracy at different conditions. An auto stack encoder has three layers: an input layer, a hidden layer, and a decoding layer [23]. The network is trained to construct its inputs again, which makes the hidden layer to learn representations of the inputs to find the accuracy, precision, and the values will be checked with the existing machine learning algorithms [10, 11, 19]. Figure 41.2 shows the Basic Auto Stack Encoder. Figure 41.3 shows the Auto Stack Encoder Training procedure.

Results Evaluation The data will be depicted based on the obtained results while using the Auto Stack Encoder will be verified based on the accuracy of the procedure. A Confusion Matrix created, shows the performance of the classification algorithm, where the future values will be predicted using the train values with the corresponding labels. Figure 41.4 shows the Genuine and Fake Classification. The results from different networks will be crossmatched using the auto stack encoder.

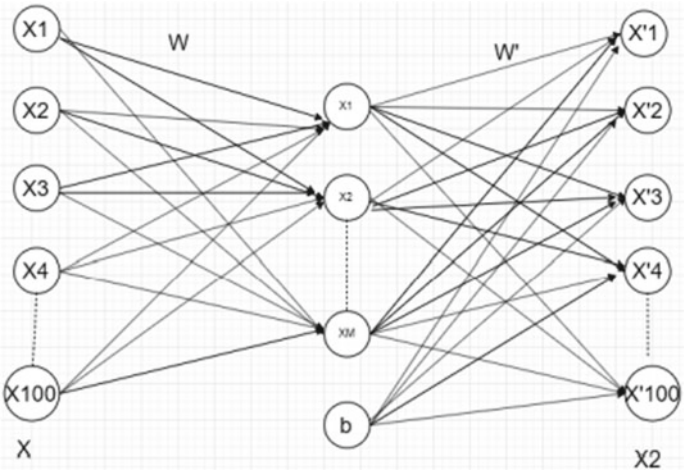


Fig. 41.2 Basic autostack encoder

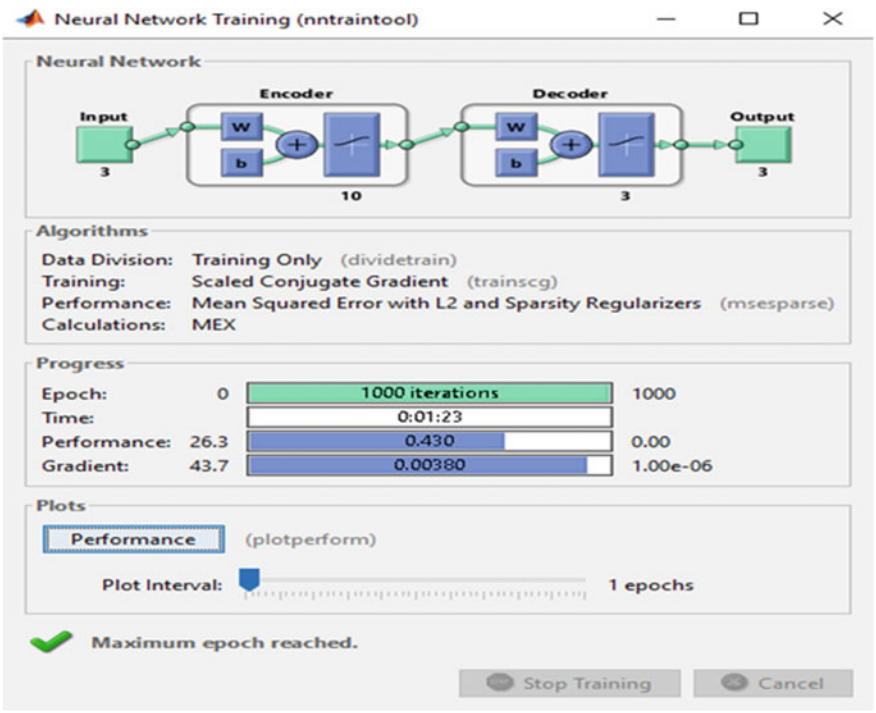


Fig. 41.3 Autostack encoder

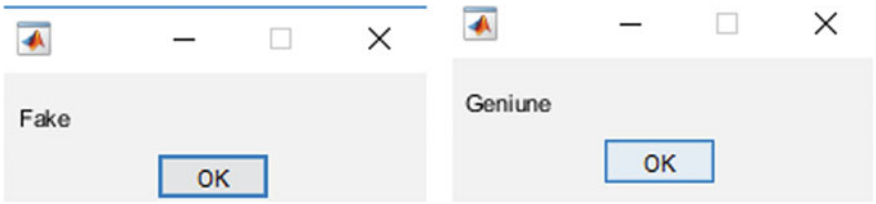


Fig. 41.4 Genuine and fake classification

41.4 Experimental Results

The proposed system was implemented using Auto Stack Encoder in the classification procedure. The accuracy of model using Auto Encoder showed the highest while others showed less. Table 41.1 shows the accuracy of the experiment using the contactless to contact-based fingerprints. Figure 41.4 shows the Genuine and Fake Fingerprint classification output. Figure 41.5 shows the graph comparing the accuracy of the RTPS + DCM Model and proposed Auto stack Encoder Model. Confusion Matrix was created using the training data with the predicted data. Here,

Table 41.1 Accuracy comparison

Database	Experiment	Accuracy (%)
PolyU Contactless to contact based fingerprint database	RTPS + DCM, *Lin et al. [1]	94.11
PolyU Contactless to contact based fingerprint database	Auto stack encoder	96

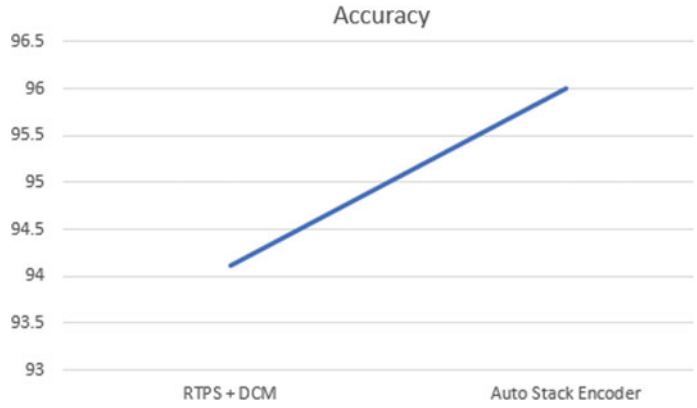


Fig. 41.5 Accuracy comparision

we obtained high accuracy for fingerprint classification compared to other existing methods.

41.5 Conclusion and Future Work

Contactless 2D fingerprint identification is more hygienic when compared to the contact-based and it enables deformation-less imaging for higher accuracy. The main function of Auto Stack encoder is for a contact-based and contactless fingerprint matching process. This trained method is using fingerprint images and extracted features from the collected fingerprints. Therefore, a fingerprint representation is formed from the combination of deep feature vectors generated from different networks. The experiment results using publicly available database illustrate the method which can achieve superior matching performance. Sensor interoperability using contact-based and contactless fingerprint is an emerging problem which is need to be addressed with other deep learning methods in the future.

41.6 Declaration

The permission from respective authorities to use the images/data as given in the paper was obtained. In case of any dispute in the future, we shall be wholly responsible.

References

1. Lin, C. and Kumar, A.: Matching contactless and contact-based conventional fingerprint images for biometrics identification. IEEE Trans. Image Process. **27**(4)
2. Bouazizi, M., Ohtsuki, T.: Contactless fingerprint identification using level zero features. In: IEEE (2016)
3. Hiew, B.Y., Teoh, A.B. and Pang, Y.H.: Touch-less fingerprint recognition system. In: IEEE (2007)
4. Jin, Q., Tago, K.: Addressing sensor interoperability problem using fingerprint segmentation. In: 2017 IEEE 10th International Conference on Service-Oriented Computing and Applications (2017)
5. Roy, A., Memon, N., Ross, A.: Masterprint: exploring the vulnerability of partial fingerprint-based authentication system. IEEE Trans. Inf. Foren. Secur. **12** (2017)
6. Bontrager, P., Roy, A., Togelius, J., Memon, N., Ross, A.: DeepMasterPrints: generating MasterPrints for dictionary attacks via latent variable evolution. In: 2018 IEEE International Conference on Biometrics: Theory, Applications and Systems (2018)
7. Tiwari, K. and Gupta, P.: A touch-less fingerphoto recognition system for mobile hand-held devices. IEEE (2016)
8. Pavithra, R., Suresh, K.V.: Fingerprint image identification for crime detection. In: 2019 International Conference on Communications and Systems (2019)
9. Ross, A., Jain, A.: Biometric sensor interoperability: a case study in fingerprints. In: Biometric Authentication, pp. 134–145. Springer, Berlin, Germany (2004)
10. Loey, M., El-Sawy, A., El-Bakry, H.: Deep learning autoencoder approach for handwritten Arabic digits recognition. In: 2017 Computer Vision and Pattern Recognition (2017)
11. Lin, C., Kumar, A.: A CNN-based framework for comparison of contactless to contact-based fingerprint. In: IEEE Transactions on Information Forensics and Security (2018)
12. Dong, G., Liao, G., Liu, H., Kuang, G.: A review of the autoencoder and its variants: a comparative perspective from target recognition in synthetic aperture radar images. IEEE Geosci. Remote Sens. Mag. (2018)
13. Szymkowski, M., Saeed, K.: A novel approach to fingerprint identification using method of sectorization. In: 2017 International Conference on Biometrics and Kansei Engineering (2017)
14. Labati, R.D., Genovese, A.: Toward unconstrained fingerprint recognition: a fully touchless 3-D system based on two views on the move. In: IEEE Transactions on Systems, Man, and Cybernetics: Systems (2015)
15. Liu, X., Pedersen, M., Charrier, C., Cheikh, F.A. and Bours, P.: An improved 3-step contactless fingerprint image enhancement approach for minutiae detection. In: IEEE (2016)
16. Alshehri, H., Hussain, M., Aboalsamh, H.A., Al Zuair, M.A.: Cross-sensor fingerprint matching method based on orientation, gradient, and gabor-HoG descriptors with score level fusion. IEEE Trans. Inf. Foren. Secur. (2018)
17. Zhai, J., Zhang, S., Chen, J., He, Q.: Autoencoder and its various variants. IEEE International Conference on Systems, Man and Cybernetics (2018)
18. Schuch, P., Schulz, S., Busch, C.: De convolutional auto encoder for enhancement of fingerprint samples. In: 6th International Conference on Image Processing Theory, Tools and Applications (2016)

19. Lin, C., Kumar, A.: Improving cross sensor interoperability for fingerprint identification. In: 23rd International Conference on Pattern Recognition (ICPR) (2016)
20. Cappelli, R., Maio, D., Maltoni, D.: Modelling plastic distortion in fingerprint images. In: Advances in Pattern Recognition—ICAPR, pp. 371–378. Springer, Berlin, Germany (2001)
21. Bazen, A.M., Gerez, S.H.: Fingerprint matching by thin-plate spline modelling of elastic deformations. Pattern Recognit. **36**(8), 1859–1867 (2003)
22. Ross, A., Dass, S., Anil, J.: A deformable model for fingerprint matching. Pattern Recognit. **38**(1), 95–103 (2005)
23. Ross, A., Nadgir, R.: A thin-plate spline calibration model for fingerprint sensor interoperability. IEEE Trans. Knowl. Data Eng. **20**(8), 1097–1110 (2008)

Chapter 42
Feature Selection by Associativity for Sentiment Analysis



S. Fouzia Sayeedunnisa, Nagaratna P. Hegde, and Khaleel Ur Rahman Khan

Abstract Sentiment analysis faces the biggest challenge of high-dimensional feature set. The current trends contemplate this intricate process through conventional feature selection/extraction. In this proposal, the statistical feature selection method, Wilcoxon sign rank score, is used for optimal feature selection. The proposal evaluates and compares the conventional filter method of feature selection with the Wilcoxon sign score for finding the optimal features. The features are defined by feature association between the terms and the sentiment lexicon. The identified optimal features using information gain (IG), chi square, and statistical method Wilcoxon sign score are then classified using the Adaboost classifier. The experiment was conducted on three datasets widely used for sentiment analysis. Result from the experiment demonstrate that the classification accuracy using the proposed method of optimal feature selection is much higher when compared to the contemporary models.

42.1 Introduction

Opinion mining is a sub-domain of text mining which deals with classification of human-generated online text as positive or negative. With the exponential growth of Internet users, accessing social media has led to generation of vast content online. The social media platforms are used to express the views, opinion about products, brands, politics, and every other field. The main focus of SA is to draw out and identify opinion from the vast online content. The opinions drawn out from this

S. F. Sayeedunnisa (✉)
M.J. College of Engineering and Technology, Hyderabad, India

N. P. Hegde
Vasavi College of Engineering, Hyderabad, India

K. U. R. Khan
ACE Engineering College, Hyderabad, India
e-mail: khaleelrkhan@aceec.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_42

online content is the major source of information which provides valuable insights [1] to the stakeholders or decision makers in improving their business [2].

Machine learning (ML) techniques for sentiment classification of user generated data includes ensemble techniques, supervised, semi-supervised, and unsupervised learning. Researchers have used lexicon-based techniques incorporating dictionary methods [3], corpus methods [4], and lexicon method with NLP [5]. Sentiment identification is a three-step process where feature extraction is the initial step, second is feature selection, and last is classification of text as in three classes, i.e., positive, negative, or neutral. Natural language processing (NLP), ML, and information retrieval techniques are used by researchers to achieve this.

The major challenge in classification of text is high dimensions of features, overlapping and redundant features [6] which takes lot of time for processing. These challenges not only affect the classifier performance but also increase the cost. The most effective techniques for dealing with these challenge are feature selection suggested by [7]. Feature selection helps in determining an ideal subset which is more proximate to the actual data.

The studies in literature [8, 9] classifies the FS methods into three categories. The first one is the filter method which calculates the score or rank of a feature. All the high rank features are selected, eliminating the low rank/score feature. In this method of FS, the relationship or association of features with each other is not considered. The second method of FS is the wrapper method which uses classifier to evaluate optimal feature subset. The last method of FS is the embedded approach which integrates feature selection into classifier. This approach faces the challenge of learning algorithm [5].

42.2 Related Work

The recent research in the field of feature selection and extraction which is predominant in selecting extremely important subset of feature called as optimal features which is referred in this section.

Research by [10] has developed proportional rough feature selector (PRFS) a unique method which is using rough set theory to identify if a term exactly belongs to a class or possibly belongs to a class. Experimental results uses Macro-F1 measure which signifies competitive performance of PRFS in view of other feature selection methods.

In the work done by [11], feature selection is done by filter methods. The methods used are term frequency (TF), information gain (IG), document frequency (DF), chi square (CHI), and mutual information (MI). These five methods are successful in eliminating 97% of the low-scoring features. The optimal feature significantly improves the accuracy.

In the work by [12], Pearson correlation-based attribute selection is used with six machine learning (ML) classifiers. The authors use string-to-word vector (STWV)

with attribute selection measure on the ML classifiers. Results demonstrate that the obtained optimal feature improves the accuracy between 78 and 88%.

In the work on feature subset selection by [13], an ordinal-based integration of different feature vectors (OIFV) is proposed as a new feature vector. In the proposed method, hybrid filter and wrapper methods are applied to obtain a frequency-based integration of optimal features as subsets. The experiment results signifies better classification accuracy.

It is evident from the recent work that integration of two methods can be more helpful toward optimal feature selection. In the proposed study, the feature association metric by [14] is used to evaluate the performance of Adaboost classifier against the filter-based feature selection techniques like IG and chi square.

42.3 Feature Selection

42.3.1 Existing Feature Selection Methods

The filter-based method of feature selection used is information gain (IG) and chi square (CHI). They are descried below:

Information Gain (IG): It is an entropy-based method for evaluating features. It is a machine learning technique used in feature selection. It is the amount of information gained [15] by presence or absence of a feature for classification.

$$\mathrm{IG}(C, A) = H(C) - H(C|A) \tag{42.1}$$

where IG(C, A) is information gain for feature A in Class C, $H(C)$ is entropy across sentiment class C, and H (C|A) is conditional entropy for feature A

$$H(C) = -\sum p(C = i) \log p(C = i) \tag{42.2}$$

$$H(C|A) = -\sum p(C = i|A) \log p(C = i|A) \tag{42.3}$$

The calculated IG values for every feature lies between 0 and 1.

Chi Square (CHI): The chi square test [16] is a statistical test to find if there is an association or correlation between variables. Chi square measures the deviation of expected counts and observed counts.

$$X^2 = \frac{\Sigma(\text{Observed}-\text{Expected})2}{\text{Expected}} \tag{42.4}$$

42.3.2 Proposed Method Using Feature Association as a Feature

Feature Association: The statistical method adapted [14] to select optimal subset of features is Wilcoxon sign rank. Feature association frequency (FAF) is computed for each feature with respect to positive and negative sentiment lexicon [17]. The association method is novel and invalidates the impact of humongous features present in social media data. The approach devised can eliminate the irrelevant features. In the starting phase, the sentiment lexicon is extracted from the labeled opinion documents. Further, the terms that appear other than lexicon are considered as features, and their FAF with respect to each extracted sentiment lexicon is calculated. The FAF is defined as the ratio of the "occurrence sentiment lexicon in association with a term and the total term occurrence."

Optimal Feature Selection using Wilcoxon Signed rank (WSR): The WSR test [18] is used to compare related samples of normally distributed data. It works on ranked data. This test is applied to the feature association frequency (FAF) of positive sentiment lexicon and the FAF of negative sentiment lexicon. We have adopted the process of optimal feature selection by [14] using WSR score. The optimal feature of positive is the one which has a greater sum of Wilcoxon positive rank when compared to the sum of negative rank.

42.3.3 Preprocessing

The training data consist of positive and negative labeled opinion records. Each row of a record set consists of bag of words. The stop words and non-textual words are removed, and further stemming is applied. This results in two sets of records *PS* and *NS* where *PS* represents all positively labeled record and *NS* represents all negatively labeled preprocessed records.

42.3.4 Sentiment Lexicon Extraction

We use sentiment lexicons provided by [17]. The features in *PS* and *NS* are compared with [17] and the extracted lexicons are stored in $l+$ and $l-$. The algorithm for feature association [14] is given below:

Algorithm Feature Association

Input : Pre- processed data labelled as positive/negative
Output : Features with Feature Association Frequency

Notation
PS: Records containing all positive labelled opinion
NS: Records containing all negative labelled opinion
l+: Set containing lexicons present in records of *PS*
l-: Set contains lexicons present in records of *NS*
wl: Unique word list existing in both *PS* and *NS* records.

Procedure Feature Selection()

```
for each pre-processed record
    for each word in review
        Search word in Opinion Lexicon[16]
        If found in Positive Lexicon then
            Insert(l+, word)
        Else If found in Negative Lexicon then
            Insert(l-, word)
        Else
            Insert(wl, word)
        Next word
        Next record
for each term wl do
    Initialize term occurrence count c(t_i) to 0
    for each record of PS do
        if the term t_i exist in record then increment c(t_i)
    End
    for each record of NS do
        if the term t_i exist in record then increment c(t_i)
    End
End
    //Sentiment lexicon and term co-occurrence(Association)//
for each term in wl do
    for each positive lexicon in l+ do
        Initialize co-occurrence counter (lexicon and term) in PS to 0
        for each record in PS do
            if term and lexicon co-occur in record
                Increment the term and lexicon co-occurrence counter
            End
        End
End // Repeat the process for negative term and lexicon
                    // FAF (Feature Association Frequency)//
for each term in wl do
    for each word in l+ which is a positive lexicon do
        Find FAF_PS as the ratio of positive sentiment lexicon and term co-occurrence
count to the term occurrence
    End
    for each word in l- which is a negative do
        Find FAF_NS as the ratio of negative sentiment lexicon and term co-occurrence
count to the term occurrence
    End
End
```

42.4 Classifier

The concept of classification in machine learning is to categorize labeled data into set of classes or groups. Predicting the target class of the test data is the main aim of classification. The classifier performs best if the training data comprises relevant observation with the class membership defined. Adaboost [19] is an ensemble classifier. It is helpful over the binary classification for boosting the performance. It is a combination of weak classifier to form strong classifier. It is an iterative method of training by retaining the algorithm.

42.5 Experimental Study

The input data is preprocessed, and feature selection techniques, i.e., IG and chi square, are applied. The optimal features of the proposed method are calculated by the WSR between the FAF of a term with positive lexicon and the FAF of the corresponding term with negative lexicons. The accuracy of the optimal features from distinctive dataset is tested on Adaboost classifier.

42.5.1 Datasets and Statistics

Experimental study of the proposed method and conventional filter methods is carried out by using three benchmark datasets, namely First GOP debate [20], Twitter sentiment analysis [21], and sentiment 140 [22]. The Twitter sentiment analysis [21] has 99,989 reviews, among them 19,998 are used and rest are excluded for analysis. The sentiment 140 [22] consists of 1.6 million of tweets, out of which 16,000 tweets were used for analysis and rest were excluded because of difficulty encountered during preprocessing (Table 42.1).

The optimal features obtained by different feature selection techniques, i.e., WSR score, IG, and CHI, are given in Table 42.2. In furtherance, classification of optimal features using Adaboost classifier is shown in Fig. 42.1

Table 42.1 Statistics of dataset

Dataset	Total reviews	Training data	Testing data	No of sentiment lexicons found in training set
First GOP debate	13,871	10,403	3467	21
Twitter sentiment analysis	19,998	14,998	4999	37
Sentiment140	16,000	12,000	4000	40

Table 42.2 Optimal features detected by feature selection strategy

Dataset	Features found in training	Optimal features by WSR	Optimal features by IG	Optimal features by CHI
First GOP debate	1535	88	175	170
Twitter sentiment analysis	2371	152	257	255
Sentiment140	2147	210	305	303

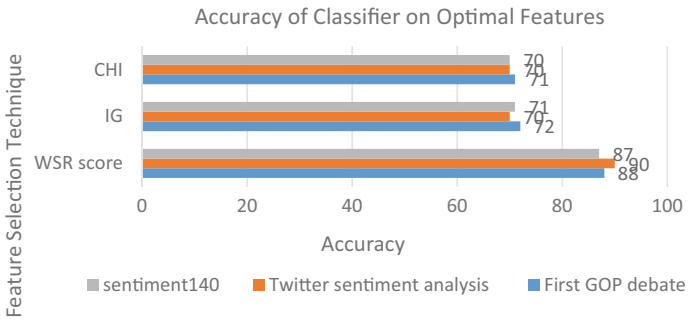


Fig. 42.1 Classifier accuracy statistics

The accuracy of the classifier for all the three dataset can be observed as shown in Fig. 42.1. Results from the experiment conducted signify that the optimal features selected by WSR test are not only less in number but also improves the classifier accuracy when compared to the filter method of feature selection. The classifier could attain a maximum accuracy of 90% on Twitter dataset using Wilcoxon Score as the feature selection technique. The key factor in improvement of classification accuracy is the relevance of features selected by Wilcoxon signed rank score.

42.6 Conclusion

In this proposal, a novel feature association technique as the sentiment lexicon singly cannot predict the class label with adequate accuracy. The feature association is the ratio of co-occurrence of a sentiment lexicon and a term to the total occurrence of the term. The optimal features are obtained by the Wilcoxon sign test of FAF. Experiment result demonstrates that the optimal features selected with Wilcoxon score gives better classifier accuracy with reduced features when compared to the filter methods. The Adaboost classifier could achieve an accuracy level around 90% across the Twitter datasets.

References

1. Wu, Y., Ren, F.: Learning sentimental influence in twitter. In: 2011 International Conference on Future Computer Sciences and Application, pp. 119–122. IEEE (2011)
2. Chung, W., Tseng, T.L.B.: Discovering business intelligence from online product reviews: a rule-induction framework. Expert Syst. Appl. **39**(15), 11870–11879 (2012)
3. Medhat, W., Hassan, A., Korashy, H.: Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J **5**(4), 1093–1113 (2014)
4. Bhatnagar, V., Goyal, M., Hussain, M.A.: A novel aspect based framework for tourism sector with improvised aspect and opinion mining algorithm. Int. J. Rough Sets Data Anal. (IJRSDA) **5**(2), 119–130 (2018)
5. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-based methods for sentiment analysis. Computational Linguistics **37**(2), 267–307 (2011)
6. Savoy, O.K.J.: Feature selection in sentiment analysis. In: Proc. CORIA No. January, 273–284 (2012)
7. Sharma, A., Dey, S.: Performance investigation of feature selection methods and sentiment lexicons for sentiment analysis. IJCA Spec. Issue Adv. Comput. Commun. Technol. HPC Appl. **3**, 15–20 (2012)
8. Ekbal, A., Saha, S.: Combining feature selection and classifier ensemble using a multiobjective simulated annealing approach: application to named entity recognition. Soft. Comput. **17**(1), 1–16 (2013)
9. Bharti, K.K., Singh, P.K.: Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering. Expert Syst. Appl. **42**(6), 3105–3114 (2015)
10. Cekik, R., Uysal, A.K.: A novel filter feature selection method using rough set for short text data. Expert Syst. Appl. **160**, 113691 (2020)
11. Rogati, M., Yang, Y.: High-performing feature selection for text classification. In: Proceedings of the 11th International Conference on Information and Knowledge Management, pp. 659–661 (2002)
12. Sharma, S., Jain, A.: An empirical evaluation of correlation based feature selection for tweet sentiment classification. In: Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies, pp. 199–208. Springer, Singapore (2020)
13. Yousefpour, A., Ibrahim, R., Hamed, H.N.A.: Ordinal-based and frequency-based integration of feature selection methods for sentiment analysis. Expert Syst. Appl. **75**, 80–93 (2017)
14. Sayeedunnisa, S.F., Hegde, N.P., Khan, K.U.R.: Wilcoxon Signed Rank Based Feature Selection for Sentiment Classification. In: Proceedings of the 2nd International Conference on Computational Intelligence and Informatics, pp. 293–310. Springer, Singapore (2018)
15. Yang, Y., Pedersen, J.O.: A comparative study on feature selection in text categorization. In Icml, vol. 97, no. 412–420, p. 35 (1997)
16. Li, Y., Luo, C., Chung, S.M.: Text clustering with feature selection by using statistical data. IEEE Trans. Knowl. Data Eng. **20**(5), 641–652 (2008)
17. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word–emotion association lexicon. Comput. Intell. **29**(3), 436–465 (2013)
18. Meek, G.E., Ozgur, C., Dunning, K.: Comparison of the t vs. Wilcoxon signed-rank test for Likert scale data and small samples. J. Mod. Appl. Stat. Methods **6**(1), 10 (2007)
19. Bloehdorn, S., Hotho, A.: Boosting for text classification with semantic features. In International Workshop on Knowledge Discovery on the Web, pp. 149–166. Springer, Berlin, Heidelberg (2004)
20. https://www.figure-eight.com/wp-content/uploads/2016/03/GOP_REL_ONLY.csv
21. https://www.kaggle.com/c/twitter-sentiment-analysis2/data
22. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224N Proj. Rep. Stanford **1**(12), 2009 (2009)

Chapter 43
Optimization of Serviceability of Mobile Communication Towers Under Capital Constraints



Pankaj Preet Singh Dhaliwal, Mahendra Kumar, and Veerendra Dakulagi

Abstract Mobile communication towers are important component of mobile communication system for civil and military applications. The major part of technical equipment of mobile communication system is concentrated on these towers. These towers have geographical separation and they are normally unmanned. In remote areas especially away from towns the alternate routes for communication in the system may not be available. Therefore high degree of serviceability of these towers is very essential. This can be ensured by having inbuilt reliability of equipment, easy fault finding, availability of repair facility, spares required for repairs and technical competence. All these factors require capital which is always a constraint. Total system is required to be optimized for desired serviceability of towers with capital constraints.

43.1 Introduction

Mobile communication in an area is provided to customers through mobile communication grid. A typical grid is shown in Fig. 43.1. The mobile communication towers are provided at various sites. Towers may be of various types [1–4].

Basic Towers: These towers are basically meant for providing communication to customers around within a specified distance. The equipment on these towers will be transmitter, receiver, antennas and power supply units (PSU). Transmitter and receiver have multi-channel facilities [5].

Control Towers: Control towers are more versatile and in addition to transmitting and receiving, they have got facility for routing and providing connectivity to many towers. They are equipped with additional controlling and rerouting circuits.

P. P. S. Dhaliwal · M. Kumar
Department of Electronics and Communications Engineering, Guru Kashi University, Talwandi Sabo, Bathinda, India

V. Dakulagi (✉)
Department of Electronics and Communication Engineering, Guru Nanak Dev Engineering College, Bidar, Karnataka 585403, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_43

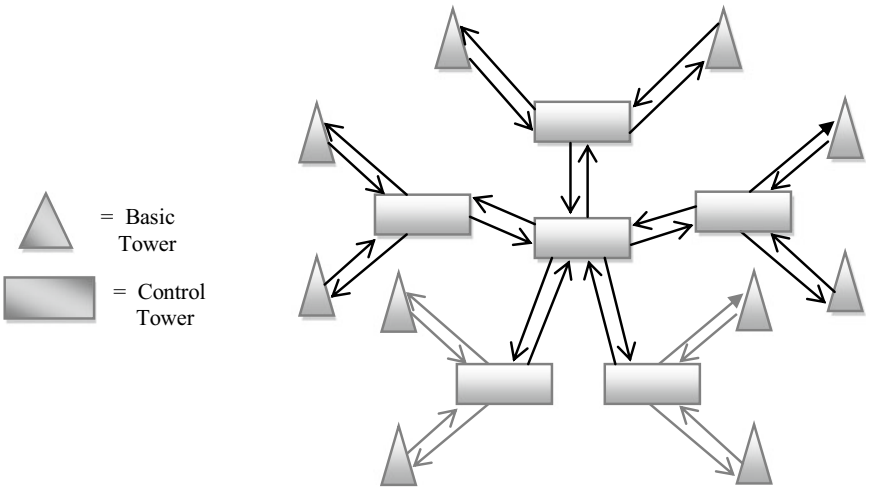


Fig. 43.1 Mobile communication grid

43.2 Points to Be Observed for Maintaining High Degree of Serviceability

Location of Towers: Although towers are generally self sustained for long periods. But there is requirement that as for as possible they should be located near to main route and roads.

Design for Maintainability: The serviceability requires certain ingredients which should be inbuilt to the system during design itself. The important factors are as under [6]:

Standardization: Components used should be minimum in range and interchangeable.

Modular Construction: The system should be designed in such a manner that the sub-systems should be in the form of modules.

Built–In-Test-Equipment (BITE): The total system should have built-in-test-equipment (BITE) which should give indication about faulty module or subsystem on front panel for easy replacement.

Defect Prone Components: The defect prone components should be easily accessible and replaceable.

Repair Facility: For modular electronic systems repair is normally restricted to replacement of modules/components with the help of built-in-test-equipment. For replacement spare parts in the form of modules/components are required to be stocked [7, 8].

43.3 Spare Parts Management

For cost effective serviceability spare parts management is very essential. The salient points are given in succeeding paragraphs.

Failure Rate of Modules/Components: Failure rate of electronic components are very less and that follow Poisson Distribution. Failure rates of modules/components should be assessed by analyzing failure data of previous years of same equipment or similar equipment.

Classification of Modules/Components: Once failure rates are established the modules/components should be classified as per following details:

Fast (F), Normal (N), Slow (S) Moving: Based on modules/components failure rate they should be classified as.

Fast Moving: If failure rate is very high.

Normal Moving: If failure rate is average.

Slow Moving: If failure rate is less than average.

Vital (V), Essential (E), Desirable (D): Based on importance of modules/components within the system, they should be classified as:

Vital: If the failure of the module/component renders the system completely unserviceable.

Essential: If the failure of the module/component renders the system partially unserviceable.

Desirables: If the failure of the module/component just reduces some desirable performance parameters.

High (H), Medium (M), Low (L) Cost: The modules/components to be classified based on their cost as under:

High Cost: Module/Components having relatively high cost.

Medium Cost: Module/Components having average cost.

Low Cost: Module/Components having less cost.

Final Classification of Modules/Components

Final Classification of Modules/Components is required to done by tagging all the three classification tags to each module/component. For example:

Module/Component ‘x’ may be Essential (E), Fast Moving (F) and Low Cost (L).

Component ‘y; may be Desirable (D). Slow Moving (S) and High Cost (H) and so on.

Lead Times: Time in days required for materializing of demands of modules/components from supplier/manufacturer are known as lead times.

Consumption of Modules/Components During Lead Times: The consumption of modules/components during lead times can be calculated from failure rates.

Example: Let failure rate of a modules/component is assessed as 3 modules/components per month and lead time is two months. Then consumption during lead time is $2 \times 3 = 6$ modules/components. Demand for components should be placed when stock position for the respective model/component is 6. So that modules/components are available in the stock when required for replacement. But this stock level of 6 will give assurance level of only 50% since failure rates of electronic modules/components are probabilistic following Poisson Distribution and stock required may be more than '6′ in about 50% cases. For achieving higher assurance levels we have to calculate modules/components required during lead time for higher assurance levels using Poisson Distribution.

43.4 Assurance Levels

Assurance Level is the probability that a module/component will be available in stock when required for replacement/repair. Desired assurance levels or availability of a particular module/component in the store for replacement will depend on its classification tags. For example a Fast Moving, Vital and Low Cost module/component should never be out of stock and therefore should have highest assurance level. A slow moving desirable and high cost modules/components should have lowest assurance level.

There are total 27 categories of modules/component as per classification (FNS, VED, and HML analysis). A typical 3—Dimensional assurance level matrix for all the 27 categories of modules/components is given in Table 43.1.

How much assurance level is required to be achieved for each category of modules/component is a management decision based on capital available for keeping stock of modules/components for repairs/replacement. A typical management deliberated 3-dimensional assurance level matrix is given in Table 43.2

43.5 Quantity of Modules/Components Required to Be Stocked for Desired Assurance Levels

Quantity of modules/components required to be stocked will depend on respective number of failures of specific module/component during lead time which is Poisson Distributed.

Table 43.1 Three dimensional assurance level matrix

Failure rate cost		L	M	H
F	V	AL_{11}	AL_{12}	AL_{13}
	E	AL_{21}	AL_{22}	AL_{23}
	D	AL_{31}	AL_{32}	AL_{33}
N	V	AL_{41}	AL_{42}	AL_{43}
	E	AL_{51}	AL_{52}	AL_{53}
	D	AL_{61}	AL_{62}	AL_{56}
S	V	AL_{71}	AL_{72}	AL_{73}
	E	AL_{81}	AL_{82}	AL_{83}
	D	AL_{91}	AL_{92}	AL_{93}

$AL_{r1} \geq AL_{r2} \geq AL_{r3}$, r = 1, 2, …, 9
$AL_{1c} \geq AL_{4c} \geq AL_{7c}$, c = 1, 2, 3
$AL_{2c} \geq AL_{5c} \geq AL_{8c}$, c = 1, 2, 3
$AL_{3c} \geq AL_{6c} \geq AL_{9c}$, c = 1, 2, 3
AL_{ab} = Assurance Level

Table 43.2 Typical 3-Dimensional assurance level matrix

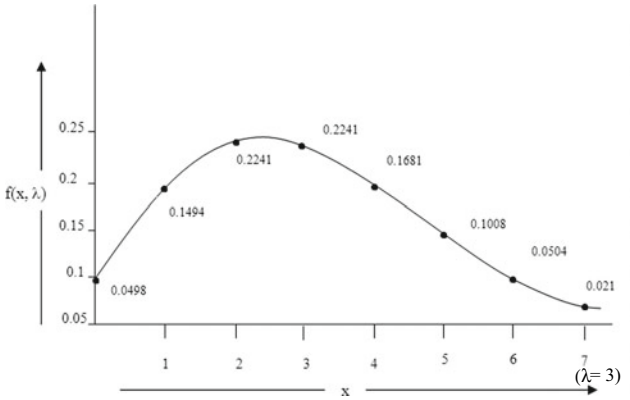
		L	M	H
F	V	80	75	70
	E	77	72	67
	D	75	70	65
N	V	72	67	62
	E	70	65	60
	D	67	62	57
S	V	65	60	55
	E	62	57	52
	D	60	55	50

Poisson Distribution:

Poisson Process:

= failure rate for modules/components during time under consideration.
T = Time under consideration.
Consider a small increment of time *t*.

(a) Probability of failure in $t = t$
(b) Probability of more than one failure $\rightarrow 0$
(c) Probability of failure of modules/components during t is independent of past happening.



x	f(x, λ)
0	0.0498
1	0.1494
2	0.2241
3	0.2241
4	0.1681
5	0.1008
6	0.0504
7	0.021

When λ = 3 (say)

Fig. 43.2 Poisson probability distribution function (for = 3)

Poisson Probability Distribution Function For = 3 (Fig. 43.2)

$$f(x, \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \tag{43.1}$$

Equation (43.1), where is λ average failure rate of modules/components.

$f(x, \lambda)$ = Probability of 'x' failure of modules/components in time under consideration.

Distribution Mean = λ Distribution Variance = λ, Distribution Standard Deviation σ (S.D) = $\sqrt{\lambda}$ Conditions: Population (n) → (large), Failure Rate per module/component = p (Very Small)

$$n \geq 20,\ p \leq 0.05$$

Possion—Commulative Distribution Function (Fig. 43.3)

$$F(k, \lambda) = \sum_{x=0}^{k} \frac{\lambda^x e^{-\lambda}}{x!} \tag{43.2}$$

$F(k,)$ = Probability of 'k' or less than 'k' failures in time under consideration.

Let the desired assurance level for a module/component (s_i) = AL_i.

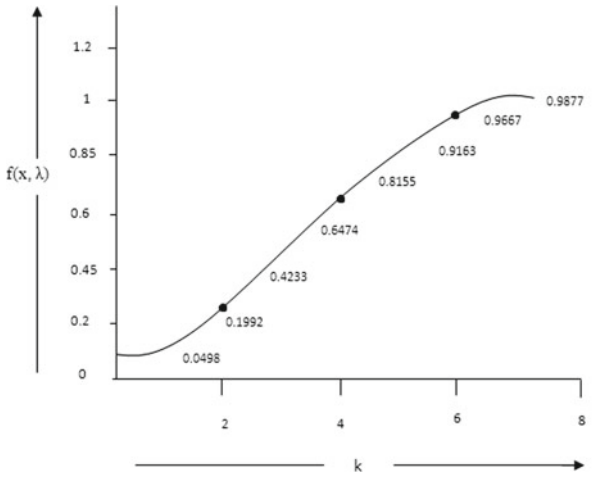
Let the failure rate/month of module/component $s_i = \lambda_i$ (Poisson Distributed).

Let the lead time of module/componentss_i in months = T_i (month).

Let No of modules/components (s_i) to be stocked for assurance $\mathrm{AL}_i = n_i$.

Then

$$\sum_{k=0}^{n_i} \frac{(\lambda_i T_i)^k e^{-\lambda_i T_i}}{k!} \geq \mathrm{AL}_i \tag{43.3}$$



k	f(X, λ)
0	0.0498
1	0.1992
2	0.4233
3	0.6474
4	0.8155
5	0.9163
6	0.9667
8	0.9877

When
λ = 3 (say)

Fig. 43.3 Poisson cumulative distribution function for $\lambda = 3$

From above equation n_i to be calculated by using Poisson Distribution given at Table 43.3.

Let total range of module/components $= N$.

$n_i =$ No of module/component (s_i) required to achieve assurance levels AL_i to be calculated for full range of modules/components ($i = 1, 2, 3, \ldots, N$).

Let the cost of module/component $s_i = c_i$.

Then

$$\sum_{i=1}^{N} c_i n_i \leq C \tag{43.4}$$

where C is total capital allotted for modules/components stocking.

Assurance levels in Table 43.2 are to be progressively adjusted/modified to satisfy Eq. (43.4) and number of modules/components required to be stocked for adjusted assurance level can be calculated by using Eq. (43.3).

43.6 Conclusion

To ensure reliability of mobile communication system, the serviceability of electronic equipment placed at communication towers is very essential. This can be achieved by having optimized design such as modular construction of technical equipment and providing of built-in-test-equipment (BITE) for identification of faulty module/component for replacements. The stock level of various

Table 43.3 Poisson distribution function

Poisson Distribution Function*

$$F(x;\lambda) = \sum_{k=0}^{x} e^{-\lambda}\frac{\lambda^k}{k!}$$

λ \ x	0	1	2	3	4	5	6	7	8	9
0.02	0.980	1.000								
0.04	0.961	0.999	1.000							
0.06	0.942	0.998	1.000							
0.08	0.923	0.997	1.000							
0.10	0.905	0.995	1.000							
0.15	0.861	0.990	0.999	1.000						
0.20	0.819	0.982	0.999	1.000						
0.25	0.779	0.974	0.998	1.000						
0.30	0.741	0.963	0.996	1.000						
0.35	0.705	0.951	0.994	1.000						
0.40	0.670	0.938	0.992	0.999	1.000					
0.45	0.638	0.925	0.989	0.999	1.000					
0.50	0.607	0.910	0.986	0.998	1.000					
0.55	0.577	0.894	0.982	0.998	1.000					
0.60	0.549	0.878	0.977	0.997	1.000					
0.65	0.522	0.861	0.972	0.996	0.999	1.000				
0.70	0.497	0.844	0.966	0.994	0.999	1.000				
0.75	0.472	0.827	0.959	0.993	0.999	1.000				
0.80	0.449	0.809	0.953	0.991	0.999	1.000				
0.85	0.427	0.791	0.945	0.989	0.998	1.000				
0.90	0.407	0.772	0.937	0.987	0.998	1.000				
0.95	0.387	0.754	0.929	0.984	0.997	1.000				
1.00	0.368	0.736	0.920	0.981	0.996	0.999	1.000			
1.1	0.333	0.699	0.900	0.974	0.995	0.999	1.000			
1.2	0.301	0.663	0.879	0.966	0.992	0.998	1.000			
1.3	0.273	0.627	0.857	0.957	0.989	0.998	1.000			
1.4	0.247	0.592	0.833	0.946	0.986	0.997	0.999	1.000		
1.5	0.223	0.558	0.809	0.934	0.981	0.996	0.999	1.000		
1.6	0.202	0.525	0.783	0.921	0.976	0.994	0.999	1.000		
1.7	0.183	0.493	0.757	0.907	0.970	0.992	0.998	1.000		
1.8	0.165	0.463	0.731	0.891	0.964	0.990	0.997	0.999	1.000	
1.9	0.150	0.434	0.704	0.875	0.956	0.987	0.997	0.999	1.000	
2.0	0.135	0.406	0.677	0.857	0.947	0.983	0.995	0.999	1.000	

*Reprinted by kind permission from E. C. Molina, *Poisson's Exponential Binomial Limit*, D. Van Nostrand Company, Inc., Princeton, N.J., 1947.

modules/components to be decided based on their failure rate, criticality and cost so that optimum stock is maintained with in capital constraints.

References

1. Amaldi, E., Capone, A., Malucelli, F., Signori, F.: Optimization models and algorithms for downlink UMTS radio planning. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'03), vol. 2, pp. 827–831 (2003)
2. Capone, A., Trubian, M.: Channel assignment problem in cellular systems: a new model and tabu search algorithm. IEEE Trans. Veh. Technol. **48**(4), 1252–1260 (1999)
3. Eisenblätter, A., Fügenschuh, A., Koch, T., Koster, A., Martin, A., Achterberg, T., Wegel, O., Wessäly, R.: Modelling feasible network configurations for UMTS. In: Anandalingam, G., Raghavan, S. (eds.) Telecommunications network design and management, pp. 1–22. Kluwer Publishers (2003)
4. Hata, M.: Empirical formula for propagation loss in land mobile radio service. IEEE Trans. Veh. Technol. **29**, 317–325 (1980)
5. Lee, C.Y., Kang, H.G.: Cell planning with capacity expansion in mobile communications: a tabu search approach. IEEE Trans. Veh. Technol. **49**(5), 1678–1690 (2000)
6. Mathar, R., Niessen, T.: Optimum positioning of base stations for cellular radio networks. Wirel. Netw. **6**(4), 421–428 (2000)

7. Mathar, R., Schmeink, M.: Optimal base station positioning and channel assignment for 3G mobile networks by integer programming. Annal. Oper. Res. **107**, 225–236 (2001)
8. Alagirisamy, M.: A new DOA algorithm for spectral estimation without source numbers information. IEEE Trans. Antennas Propag. **68**(09), 6675–6682 (2020)

Chapter 44
KruthVidya—Mobile App to Detect Dyslexia



Apurva Dasari, Rashmitha Proddaturi, Sai Prasanna Vasaraju, and Kalyani Nara

Abstract Dyslexia is one such learning disorder that makes it difficult for people to master the skills to read, write, and spell out words. This paper introduces a mobile app KruthVidya that helps to diagnose the prevalence of dyslexia by tracking the child's ability in reading and writing spellings and the coordination between these activities simultaneously. Mobile app is developed to capture the spoken utterances of a given word and written activity performed simultaneously and process this using Google Speech APIs and optical character recognition (OCR) to extract the character sequences. The results from these modules are given to classifier, which process the information to diagnose the prevalence of dyslexia using standard metrics. In general, this tool avoids the stress which parents and kids undergo during the assessment, for its design in user friendliness. It is a play app that can work similar to the standard tools used by the clinical psychologists.

44.1 Introduction

The first 5–7 years of age is very important and plays a crucial role in the rest of life of an individual. It is considered as the fastest period of growth where the brain develops in a sensitive way which is interpreted by various milestones exhibited by the child. This period of child growth is considered as the foundation for subsequent attainment of skills and knowledge at individual level which has an impact on the human resource at population level and on economic development at large. Developmental disabilities may cause impairments of a child's physical, learning, or behavioral functioning. These may be categorized as sensory impairments (hearing or vision loss), ASD—autism spectrum disorder, ADHD—attention deficit hyperactivity disorder epilepsy or seizures, cerebral palsy, intellectual disability or learning disorder. The Lancet

A. Dasari (✉) · R. Proddaturi · S. P. Vasaraju · K. Nara
G Narayanamma Institute of Technology and Science, Hyderabad, India

K. Nara
e-mail: nara.kalyani@gnits.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_44

series since 2007 provides estimates relating to early childhood development and risk of suboptimal development in low-income and middle-income countries [1]. The Global Burden of Diseases, Injuries, and Risk Factors Study 2016 has given few comprehensive and comparable estimates of age-specific health disorders for 195 countries from 1990 to 2016. Children younger than 5 years are at suboptimal development in low-income and middle-income countries.

The most common neurodevelopmental disorder effecting 3%-10% of children is specific learning disorder. Though it can be diagnosed at early stage, it is unnoticed until the adolescent age due to lack of awareness among parents and teachers.

44.2 Epidemics of Dyslexia

Good research on source and cause for dyslexia is published. It is observed that there is significant relation between reading disability and psychopathology. It was proved in the study that individuals with reading disability have displayed significantly higher rates of all internalizing and externalizing disorders than individuals without this disability [2]. Socioeconomic status (SES) is associated with reading skills. Lower SES was associated with both poorer word reading and poorer reading comprehension, and this was the observation made by MacDonald in the study published in [3]. This is because the children from lower SES families that undergo diagnostic of developmental dyslexia would be less.

Initially, the research on dyslexia focused primarily on the reading difficulties in English, while the recent study is focused on the nature of dyslexia across various languages. To summarize on what is known by this research, dyslexia manifests across languages exhibiting two types of variety, first, among alphabetic orthographies that vary in the degree of consistency of letter-sound correspondences, and second, in alphabetic versus logographic orthographies. Children at the low end of reading ability distribution in languages with more consistent mappings between letters and sounds (e.g., Italian or Finnish) have less severe reading problems than those learning to read less consistent languages (i.e., English), in the comparison study made with English and German dyslexics showed the evidence of systematic differences in the reading performance due to the difference in orthographic consistency which was similar for normal and dyslexic children [4]. Difficulties with reading fluency (speed of reading connected text) seem similar across languages [5]. Several studies have noted important universal features in normal and disordered reading across cultures despite linguistic differences. Cognitive predictors of early reading were observed to be similar for five European orthographies (Finnish, Hungarian, Dutch, Portuguese, and French).

44.3 Factors in Phonological Theory and Their Relation to Dyslexia

44.3.1 Access Versus Representations

Since the ability to distinguish phonemes was innate and universal, the implicit phonological representations should be intact in dyslexia. Impressive amount of evidence is available to indicate that a huge number of dyslexics have a phonological deficit that is the ability to access, process, and manipulate speech sounds and written text units [6, 7]. Two dominant theoretical frameworks were proposed to explain the phonological deficits in dyslexia. Classical phonological theory of dyslexia postulates the phonological difficulties as the core deficit of dyslexia. According to this postulate, underspecified phonological representation results in specific processing difficulties that are phonological in nature. Alternative theories are proposed explaining that phonological difficulties are due to indirect impact of sensorimotor impairments in the auditory system [8], visual system [9], and the cerebellum [10].

It appears that for all children, initial learning starts with phonological representations as fairly holistic and become gradually increased over time. Kids represent most words as single entity in early childhood. Through language development, phonological representations begin to emphasize syllables, then sub-syllabic distinctions, and ultimately individual phonemes. In his contributions argued that this implicit phonological development underlies parallel development in metaphonological awareness. In fact, in preschool children, they generally cannot perform tasks that require explicit manipulation of individual phonemes, although they can perform phonological awareness tasks at a larger grain size (e.g., rhyming). Children with dyslexia show deficits on phonological processing tasks like manipulation of phonemes, speech perception, priming, and lexical gating. Various research groups have described that the implicit phonological deficit may be observed in different ways by the psychologists.

44.3.2 Causal Direction

The second factor in the phonological theory of dyslexia is concerned with the direction of effect between phonological development and reading [11]. Written language is parasitic on oral language. In general, children do not begin writing until they have mastered most of the fundamentals in spoken language, and it seems reasonable that the causal direction flows from phonology to reading rather than vice versa. Several observations and evidence support this conclusion. Children with dyslexia underperform even younger, typically on reading level on phoneme awareness tasks [12]. The study proposed by [13] showed that these deficits tend to persist in adults with dyslexia who have otherwise compensated well for the disorder. These studies conclude that phoneme awareness deficits have a unidirectional causal link

to reading problems which is oversimplified for several reasons. However, there are long standing controversies about the units of speech perception, and few evidences demonstrate that speech representations preserve much more than phonemes [14]. Adults natural illiterates, who are cognitively normal but have no formal schooling, confirm that phoneme-level representation is not important in language development. Difficulties in phonological development in dyslexia are probably not restricted to phonemic or segmental representations and must lie in other dimensions of the speech stream.

44.3.3 Orthographic Learning

Orthographic learning emphasis to establish mapping between phonemes and graphemes. Good neurophysiological evidence indicates that skilled readers treat letters as single audio-visual objects, and the orthographic learning hypothesis states that problems developing such integrated representations interfere with the emergence of fluent reading. Limited studies are reported to support this hypothesis, and research is carried out in comparing the performance of children with and without dyslexia, training them on associations between sounds in their native language and an unfamiliar orthography. Results indicate that the children with dyslexia performed more poorly, under time pressure. The orthographic learning has an advantage of integrating the brain and neuropsychological levels of analysis to explain reading, development, and difficulties. The limitation is that it is difficult to test a pure integration account of phoneme-grapheme binding as the children with dyslexia are not equivalent to their typically developing peers in processing phonemes of their native language.

44.4 Motivation

With one child in around every 250 children affected with developmental disorder (DD), India stands fourth place worldwide. Country has nearly 1.3 billion people with children $\leq$ 15 years constituting nearly one-third of the population. An estimate from 2019 states that more than two million people might be affected with disorders. Tools are available for assessment of children above 7 year and are used only when the child is taken to psychologist for assessment. Dyslexia is not a disease. It is a condition a person is born with, and it often runs in families. It is estimated that every 1 in 10 people have dyslexia. Research has shown that dyslexia happens because of the way the brain processes information. Most dyslexic people have average or above-average intelligence, and they work very hard to overcome their learning problems. Pictures of the brain show that when people with dyslexia read, they use different parts of the brain than people without dyslexia.

People with dyslexia often have normal intelligence and vision. Dyslexics may face problems in spelling words, reading quickly, writing words, "sounding out" words in the head, pronouncing words when reading aloud, and understanding what one reads. Early signs such as late talking, learning new words slowly, problems forming words correctly, problems remembering or naming letters, etc., could often be misleading suggesting other disabilities. Hence, when a child is in school, dyslexic signs and symptoms become more significant such as problems processing and understanding what he or she hears, problems remembering sequence of things, difficulty seeing similarities and differences in letters and words, difficulty in spelling words, etc. It is not unusual to identify signs in teens or adults. A teen's parents or teachers can identify signs such as poor reading and writing skills, despite normal intelligence, poor spellings, difficulty remembering names and lists, problems with directions, etc. Having these signs necessarily does not mean dyslexia. It can be confirmed by standardized tests under supervision.

As mentioned, people with dyslexia are often poor spellers and find it difficult in segmenting words into individual sounds or may blend sounds when producing words. They may also exhibit signs of difficulty in identifying or generating rhyming words or count the number of syllables in words. Hence, considering the writing of spelling of the word for diagnosis of dyslexia is a good measure. Dyslexic people not only find it difficult to write the words but also find it difficult to spell them out. They find difficulty in spelling out longer words, or remembering names, letters, numbers, etc., sounding out little words. The second element that this application considered is the reading out the word given for diagnosis. It is a good measure to see if the user is able to identify the correct letters and phonetics of the word provided. The final understanding is the coordination of the speech and written input which identifies the brain's activity while performing the two activities simultaneously. Normal people can perform the reading and writing simultaneously while dyslexic people find it difficult to coordinate. The early symptoms are unnoticed for two reasons. 1. Indian parents or care takers ignore the early symptoms due to lack of awareness in observing the exhibited skills at an early stage. 2. The teachers in pre-primary schools are not trained, and most of the schools do not have a clinical psychologist appointed to perform primary screening. These challenges can be addressed with KruthVidya a mobile tool that can assist both parents or care takers and pre-primary teachers to diagnose the prevalence of dyslexia and refer to clinical psychologist to facilitate early intervention.

In Indian context when the child is observed to underperform, they are taken to clinical psychologist for assessment. There are standards defined by National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru. Few examples of tests that are currently in use as defined by NIMHANS are mentioned below:

- The NIMHANS test has two stages of assessment: Assessment one for pre-academic skills; attention, visual, and auditory discrimination, visual and auditory memory, speech and language, writing and number skills. Speech and language, writing and number skills. Level 2 test areas of attention, reading, spelling, perceptuo-motor, visuo-motor integration, memory, and arithmetic skills. It is

used for assessment of SLD and used to monitor progress after remediation. A wide range of relevant areas are covered. It is easily administered and available. It is good for early identification of learning issues. It can only be assessed by clinical psychologists and special educators.

- WRAT-4 (Wide Range Achievement Test): The tests included are reading composite (word reading + reading comprehension), spelling, and math computation. It can be assessed only by psychologist, special educators, and teachers. The advantages include screening of difficulties in academic areas, take 15–30 min, and have parallel forms allowing for retesting, age- and grade-based norms.
- TOWL-4 (Test of Written Language): The tests included are vocabulary, spelling, style (punctuation and grammar), logical sentences, sentence combining, contextual conventions, contextual language, and story construction. This test can be assessed by anyone with formal training in assessment and English proficiency. The advantages would help in identifying students with writing difficulties, strong conceptual model of writing, good reliability, excellent tool to assess the progress of remediation in writing. The disadvantages would be that it is time consuming, subjectivity in scoring procedures in domains, and it cannot be used to assess other areas of SLD except writing.

The above tests are currently under use and can be assessed only by psychologists, trained professionals, and teachers. This model "KruthVidya" is a mobile app to assist parents or guardian to identify the prevalence of dyslexia at earlier stages. The application KruthVidya allows children to interact with the app as a learning tool and parallelly process the activities of the child in learning to classify into being dyslexic or not. KruthVidya is a mobile application that can detect dyslexia at an early stage. It involves evaluation of writing and phonological skills, along with coordination between child's hand and brain. The architecture of the model is explained in subsequent section.

44.5 Architecture of KruthVidya

This application has a suite of useful assistive technology features for kids with dyslexia. It is a prediction software and is developed for dyslexic children to help them in learning, reading, and recognizing. It promotes learning in an interactive manner and in an innovative learning environment. KruthVidya mobile application is proposed that would test the reading and writing skills of children, and the inability of exhibiting those skills is recognized. The application would suggest a word and ask the user to spell and write the word simultaneously. The accuracy of the activities is tested, and disorder is identified based on various factors.

The working of the application is provided words that are asked to spell by children which is recorded and converted into text in the backend and parallelly the spelt word is asked to write. The pressure of writing and the pattern, behavior in which text is written, is noted, and spelling is checked. The way that the words are pronounced

is compared to the way that the words are written. It has five development phases—analysis, design, development, implementation, and evaluation.

The evaluation is done at the end, and dyslexic characteristics are identified at an early stage. It is built to encourage independent reading for kids with learning disorders.

The application is implemented in a four-stage model and three modules that perform different activities of the implementation (Fig. 44.1). The four stages are given as follows:

a. Collection of inputs-audio and written
b. Optical character recognition implementation on the text collected.
c. Speech-to-text conversion on the audio input collected.
d. Analysis and results.

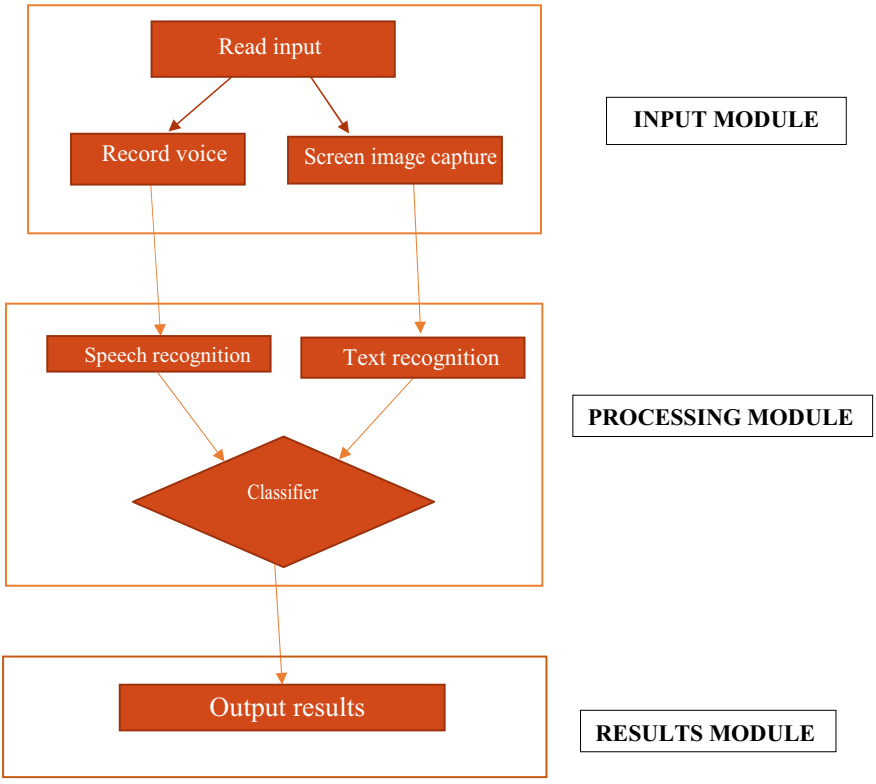


Fig. 44.1 Architecture of KruthVidya

44.5.1 Collection of Inputs

In this stage, the application provides an interface that provides the image or the name of the word that is considered for analysis. The interface has a canvas where the user is asked to write the word, and the microphone symbol enables to record the audio input of the word described. The inputs are stored to be sent into the next stages where they are analyzed. The words that are considered for analysis are in a random order initially and later move on to appear in such a way that the words or phonetics that are spelt wrong too often are given to the user. For example, if bat is the word initially and the user writes the 'b' as 'd', the model gives another word starting with 'b' say 'ball' so as to check if the user is actually finding it difficult with the letter 'b'. This process is the input module.

44.5.2 Implementation of OCR to Extract Text

The written input on the canvas from the previous stage is recognized here. This process is the processing module.

Optical character recognition (OCR) technology is used to convert the written text to digital text form. The user is allowed to use notepad for free hand writing. This text is scanned and processed by OCR to extract the digitally coded text. The process involves preprocessing the scanned imaged and compares with multiple formats to extract each individual alphabet (Fig. 44.2). Most fonts are now easily available and support a variety of digital image file format inputs, and so higher advanced systems are capable of producing high degree of recognition accuracy. It takes several steps to convert an image file into an editable document. After scanning the image, it is stored in a TIF format that is a bit-mapped file.

In this processing module, after receiving the input image from screen, the images are cropped to document segments. Then, detecting and identifying the words/alphabets in each segment and preprocessing character images then using optical character reader, the character sequence is identified. The inferred character sequence is concatenated to form the whole word, and then, the text in higher level either as word or phrase is outputted.

Using this technique, the computer tries to recognize the entire characters and matches it to the matrix of characters stored during speech-to-text conversion. In order to recognize text effectively, the software must preprocess the images.

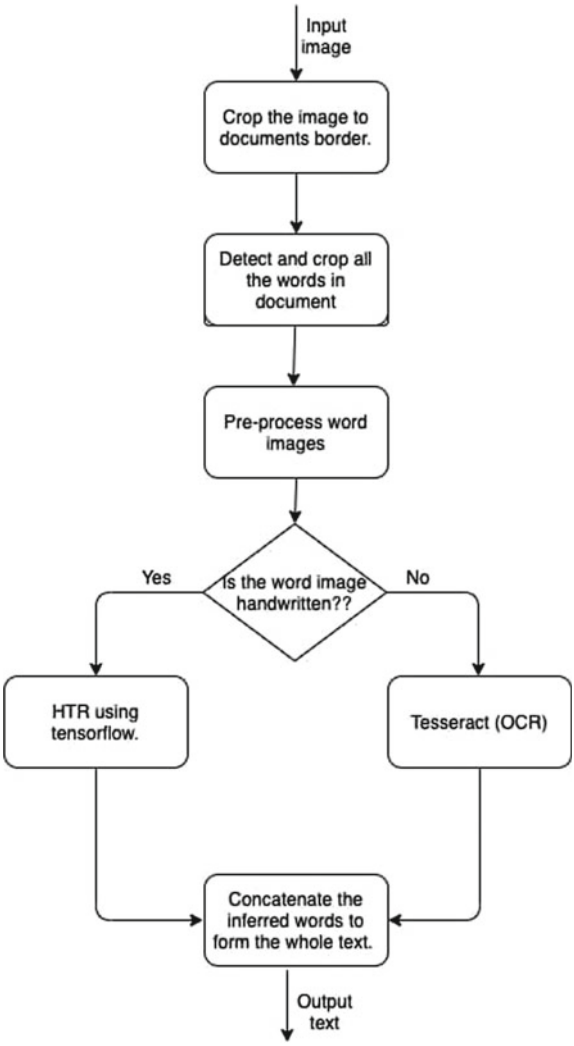


Fig. 44.2 Working of OCR model

44.5.3 Speech-To-Text Conversion of the Audio Input Collected

The audio input from the activity is used to convert into text. This process also lies in the processing module.

In the application, Google Speech-to-Text API is used to perform this operation. Speech-to-text has three main methods for recognizing speech.

1. **Synchronous Recognition**: It sends audio data to the speech-to-text API, performs recognition on that data, and returns results after all audio has been

processed. Synchronous recognition requests are limited to audio data of 1 min or less in duration.

2. **Asynchronous Recognition**: This method sends audio data to the speech-to-text API and initiates a long-running operations. For audio data of any duration up to 480 min, asynchronous requests can be used.
3. **Streaming Recognition**: It performs recognition on audio data provided within a gRPC bi-directional stream. Streaming requests are designed for real-time recognition purposes, such as capturing live audio from a microphone. Streaming recognition provides interim results while audio is being captured, allowing result to appear.

A request is sent, which identifies speech and returns the text as the response.

Google Speech-to-Text also performs audio encoding. It refers to the manner in which data is stored and transmitted. Several techniques like FLAC, LINEAR16, and OGG_OPUS are used for encoding the audio data. FLAC, however, is both a file format and an encoding. Within the speech-to-text API, FLAC is the only encoding that requires audio data to include a header; all other audio encodings specify header-less data.

The supported audio encodings are:

Codec	Name	Lossless	Usage
MP3	MPEG audio Layer III	No	MP3 encoding is beta feature, only available in v1p1beta1
FLAC	Free lossless audio Codec	Yes	16-bit or 24-bit requires for streams
LINEAR16	Linear PCM	Yes	16-bit linear PCM encoding
MULAW	μ-law	No	8-bit PCM encoding
AMR	Adaptive multi-rate Narrow band	No	Sample rate must be 8000 Hz
AMR_WB	Adaptive multi-rate wide band	No	Sample rate must be 16,000 Hz
OGG_OPUS	Opus encoded audio Frames in and Ogg container	No	Sample rate must be one of 8000, 12,000, 16,000, 24,000, and 48,000 Hz
SPEEX_WITH_ HEADER_BYTE	Speex wideband	No	Sample rate must be 16000 Hz

This API also used speech adaption to improve the transcription results. For example, suppose that your audio data often includes the word “weather.” When speech-to-text encounters the word “weather,” you want it to transcribe the word as “weather” more often than “whether.” In this case, you might use speech adaptation to bias speech-to-text toward recognizing “weather.” Speech adaption increases the accuracy of frequently used words and phrases and improves the accuracy of speech transcription when the supplied audio contains noise or is not very clear.

44.5.4 Output Results

The two inputs, i.e., the audio and written inputs, are now converted into normal text which can be compared and checked for accuracy. The primary analysis performed here verifies if the audio input is wrong or the written input has an error (Fig. 44.3). The next analysis is to identify the particular sound, alphabet, or phonetic sound

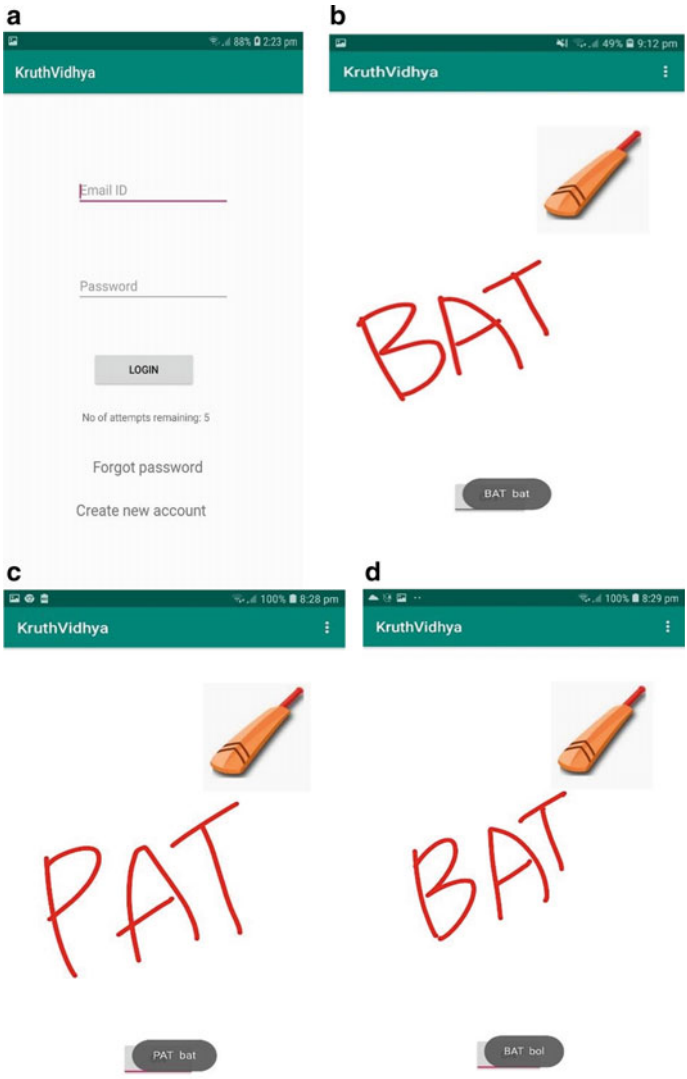


Fig. 44.3 **a** Login page. **b** Detection of correct written text and correct utterance. **c** Detection of wrong spelling and correct utterance. **d** Detection of correct spelling and wrong utterance

that the user is finding difficulty in. To confirm the same, the child is provided with words containing those phonetics or alphabets, and performance is validated. The application is currently under development implementing the simultaneous working of audio reading and the OCR technology running on the text written on the interactive surface. The current area of work is the database that is used to save scores and understand the phonetic sounds where the child or user is facing a problem with.

44.6 Conclusion

The application is a very user-friendly tool which facilitates the user to use it more like a play tool but simultaneously identifies the symptoms of dyslexia. The user would be able to use the application in a relaxed manner like a game rather than a test which is an added advantage as there is no stress and results could be more accurate. Early identification and intervention with kids who show the warning signs of dyslexia are critically important for better outcomes later on. Dyslexia is described as a hidden disability. It affects a child's ability to recognize and manipulate the sounds in language. Kids struggle with deficits in working memory that make it difficult to retain and use information in the short term. Working memory deficits impact phonological awareness, decoding, and fluency.

So, using an application technology which is suitable for kids might change the learning environment. The main objective of the KruthVidya application is to identify the learning disabilities at an early stage in an easy manner and also in an interactive manner with the kids.

References

1. Herrman, H., Swartz, L.: Promotion of mental health in poorly resourced countries. ScienceDirect © 2007 Elsevier Ltd Published, 04 Sep 2007
2. Willcutt, E.G., Pennington, B.F.: Psychiatric comorbidityin children and adolescents with reading disability. J. Child Psychol. Psychiatry **41**, 1039–1048 (2000)
3. National Center for Education Statistics (NCES). The Nation's Report Card : Reading 2011. NCES2012–457. Washington, DC: US Department of Education, Institute of Education Sciences (2011)
4. Landerl, K., Wimmer, H., Frith, U.: The impact of orthographic consistency on dyslexia: a German—English comparison. Cognition **63**, 315–334 (1997)
5. Caravolas, M., Volín, J., Hulme, C.: Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: evidence from Czech and English children. J. Exp. Child Psychol. **92**, 107–139 (2005)
6. Ramus, F.: Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? Curr. Opin. Neurobiol. **13**, 212–218 (2003)
7. Snowling, M.J., Hulme, C.: Evidence-based interventions for reading and language difficulties: creating a virtuous circle. Br. J. Educ. Psychol. **81**, 1–23 (2011)
8. Tallal, P.: Auditory temporal perception, phonics, and reading disabilities in children. Brain Lang. **9**(2), 182–198 (1980)

9. Livingstone, M.S., Rosen, G.D., Drislane, F.W., Galaburda, A.M.: Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. Proc. Nat. Acad. Sci. USA **88**, 7943–7947 (1991)
10. Nicolson, R.I., Fawcett, A.J., Dean, P.: Dyslexia, development and the cerebellum. Trends Neurosci **24**, 515–516 (2001)
11. Castles, A., Wilson, K., Coltheart, M.: Early orthographic influences on phonemic awareness tasks: evidence from a preschool training study. J. Exp. Child Psychol. **108**, 203–210 (2011)
12. Wagner, R.K., Torgesen, J.K.: The nature of phonological processing and its causal role in the acquisition of reading skills. Psychol. Bull. **101**, 192–212 (1987)
13. Doi: https://doi.org/10.1016/S0140-6736(07)61244-6
14. https://towardsdatascience.com/how-to-use-google-speech-to-text-api-to-transcribe-long-audio-files- 1c886f4eb3e9
15. https://cloud.google.com/speech-to-text/docs/concepts
16. Daud, S.M., Abas, H.: Conference paper: 'Dyslexia Baca' mobile app-the learning ecosystem for dyslexic children
17. https://www.researchgate.net/publication/269032733_%27Dyslexia_Baca%27_Mobile_App_--_The_Learning_Ecosystem_for_Dyslexic_Children

Chapter 45
A Survey on Streaming Adaptation Techniques for QoS and QoE in Real-Time Video Streaming



Satyanarayana Reddy Marri and P. Chenna Reddy

Abstract Video streaming is a dominant application in our day-to-day life and is accounting for the majority of the traffic over the Internet. HTTP streaming with dynamic adaptation (DASH) is accommodating prominent streaming services such as YouTube, Netflix, Amazon Video, and Hulu. It is a challenging task to stream the video up to the end-user's level of expectation for the service providers due to the sudden fluctuations in bandwidth availability and many more factors over the network. At present, video streaming is rated based on network-centric (QoS—quality of service) and user-centric (QoE—quality of experience) parameters. The perception of the viewer's video quality is measured by QoE metrics. In this contemporary survey, we manifest solely state-of-the-art bitrate dynamic adaptation algorithms for streaming over HTTP. The main contribution of this paper is focusing on video coding standards, diversified bitrate adaptation techniques, network-centric parameters like bandwidth, throughput, packet loss, delay or jitter which account for the QoS, QoE subjective and objective metrics, and influencing factors like initial delay, stalling rates, and quality oscillations which are taken into account. We also consider the factors for bitrate adaptation techniques like bandwidth estimation, playback buffer status, device configuration, content type, and viewer expectations. We make a note of machine learning (ML) algorithms which play a vital role in real-time data streaming. This comprehensive survey is useful to the researchers to predict the suitable bitrate chunk by the adaptation technique according to the network conditions as well as take the client-side factors into an account to optimize the viewer's QoE in real-time multimedia streaming.

S. R. Marri (✉) · P. C. Reddy
Jntua CEA Anantapur University, Ananthapuramu, India

P. C. Reddy
e-mail: chennareddy.cse@jntua.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_45

45.1 Introduction

Video streaming over the network is emerging as a ubiquitous application, and it constitutes a major portion of the overall Internet traffic. With the advent of a variety of handheld devices capable of wirelessly connecting to the Internet, the widespread use of real-time data streams has rapidly increased [1]. This type of traffic is expected to grow exponentially in the coming years. In the past few years, data streaming applications are leading IP networks and are expanding. The forecast analysis of the Cisco Visual Networking Index reveals that video traffic alone will be 82% of the global network traffic by 2021, which is 73% in 2016 [2]. It is a challenging task for content, network, and service providers to meet the user's anticipated video services. Therefore, in current real-time streaming multimedia applications, a novel solution is needed to maintain the demand for high-rate media streams, so that throughput is stable for a longer period and also to achieve high QoS and satisfactory QoE from the users. To meet this demand, we are using streaming techniques over HTTP.

Traditional video streaming (TVS) is performed based on a push-based mechanism. TVS uses protocols like real-time transport protocol (RTP) [3] and real-time messaging protocol (RTMP) [4]. To establish the session for video streaming and identify the track of the information during this period, real-time streaming protocol (RTSP) [5] is needed. The traditional communication with UDP protocol has some issues considering the perspectives of network address translation (NAT) and firewalls [6]. In current real-time streaming multimedia applications, a novel solution is needed to maintain the demand for high-rate media streams, so that throughput is stable for a longer period and also to achieve higher QoS and satisfactory QoE from the users.

In the challenging situation where both QoS and QoE are faced with multiple limitations of delivery requirements, machine learning (ML) technology can provide a promising and elegant solution, because ML can sense, mine, predict, and infer system parameters [7, 8] and can propose the best adaptive conditions as needed. This research effort aims to explore and propose a novel solution to improve QoS and QoE in real-time data streams using ML technology in heterogeneous networks. The remaining paper is manifested as Sect. 45.2 explores motivation to choose this area, streaming techniques, as well as coding standards which are explained in Sect. 45.3, and Sect. 45.4 gives solely information regarding QoS and QoE in real-time data streaming. Section 45.5 presents the role of machine learning in streaming, Sect. 45.6 gives comprehensive information on streaming adaptation algorithms in QoS and QoE, and finally, Sect. 45.7 gives the conclusion.

45.2 Motivation

The Internet is a heterogeneous environment that connects various network technologies [9]. Even as networks are supported by service categories, the network

resources available to different broadcast applications will change dynamically over time, due to requirements of network applications changes periodically [10, 11]. It is necessary to analyze a large number of real-time streaming data of complex nature to effectively meet the QoS requirements of services. It is planned to deliver through it, although it changes over time and the workload and network conditions are different [12, 13]. For humans, it is very difficult to process the bulk data in a small span to operate network services. So, it is mandated to go for automation of these tasks.

Even the widely used TCP protocols [14, 15] on the Internet use implicit feedback through timeouts and duplicate acknowledgments of lost packets. During severe congestion, the loss rate is high, which occurs when data users compete for scarce network bandwidth in the data network. Internet measurements indicate that the demand for bandwidth-intensive applications is growing, leading to increased loss rates on various links on the Internet [16]. End nodes typically deploy explicit feedback. However, implicit or explicit feedback relying solely on end nodes is not sufficient to achieve high throughput on the Internet. Appropriate mechanisms should be used to control network congestion. Otherwise, the network may end up in a persistent overload state, which may cause congestion collapse and affect the QoS and QoE of the client [7, 8]. Thus, this versatility and complexity of real-time data streams have led to the development of efficient algorithms for throughput improvement in heterogeneous environments by using machine learning techniques to reduce the amount of packet loss by making reasonable use of network resources.

45.3 Overview of Adaptive Streaming Techniques and Video Coding Standards

At present, over the top (OTT) services like Netflix, Amazon, and Hulu using the streaming algorithms like HAS and DASH to provide QoS and QoE to the end-users in dynamic fluctuations in bandwidth availability in heterogeneous networks as well as devices [14, 17]. In HTTP adaptive streaming (HAS), the video is segmented into chunks with equal size in a fixed duration of play, and each chunk is available in different bitrates [18]. The client sends an HTTP request to the streaming server, and an adaptation algorithm identifies the suitable bitrate chunk to the network conditions and transmitted it to the client using transport protocols of either TCP or UDP. The adaptive streaming process is shown in Fig. 45.1. The adaptation logic is implemented either client side or server side; generally most of the cases, it is a client side to reduce the burden to streaming server. Whenever network condition changes, then the adaptation logic requests the appropriate bitrate chunk to the client system for streaming the video seamlessly. This adaptability of HAS is being analyzed on a single client and shows impressive results [19].

For managing the exponential growth of video traffic over a public network, [14] proposes a new streaming technique over HTTP with dynamic adaptation (DASH) nature according to the network conditions. Here also, as HAS video data is split

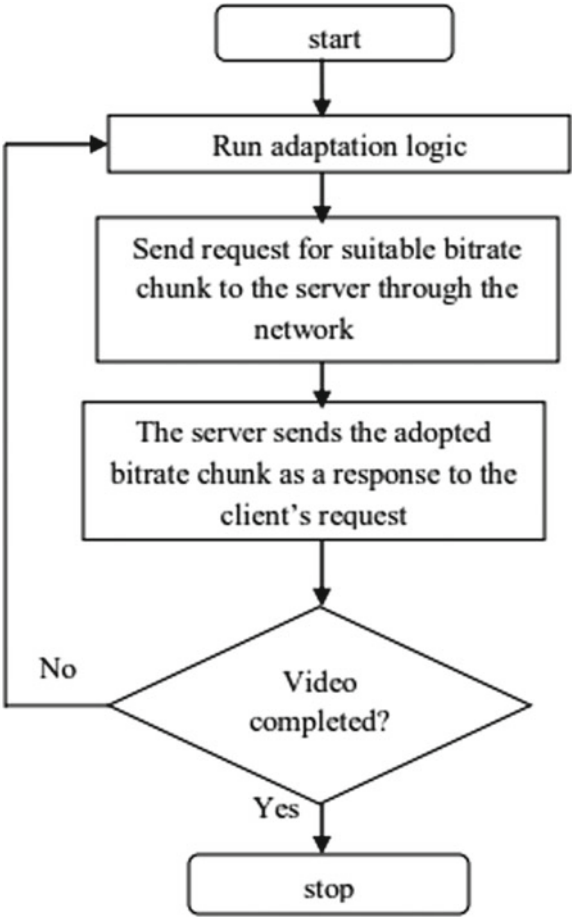


Fig. 45.1 Adaptive streaming process

into chunks with equal playback length, but each chunk is available in different bitrates as well as quality levels [11]. The target of DASH is also to provide seamless streaming of data to the user's expected level. A dynamic adaptive technique that enhances HTTP adaptive streaming (HAS) [14, 18] is a prominent solution in video streaming, and it needs new mechanisms for flow control, free from congestion, and efficient usage of available bandwidth. Video coding standards are the compression techniques that play a vital role in transmitting the data very fast over the Internet. That means the data is compressed before sending it and decompressed at the destination. In video streaming also, first, the video is segmented into chunks which are equal in size, and later each chunk is compressed (encoded) into different bitrates. Several video coding standards differ from each other in their performance and efficiency like H.264 [20], SVC [21], HEVC [22], and SHVC [23].

45.4 QoS and QoE in Real-Time Data Streaming

In recent years, most of the research work move around QoS and QoE due to video traffic currently leading the Internet [24, 25]. Most of these indicate that QoS assessment has shifted from network performance to QoE. QoE measures the performance of a user's subjective perception. However, QoE and QoS are not mutually exclusive, but QoE is an extension of QoS. While we are developing the QoS model, it is mandatory to consider all QoS parameters of the network and more importantly to quantify the relationship between them. The QoS framework of real-time data flow [26] faces a huge success challenge. This can be a daunting task when dealing with next-generation heterogeneous network systems [27, 28], where many unpredictable instances can affect the user experience. The work in [29] proposes a model that guarantees for handling delay-related sensitive applications. The algorithm of [30] identifies the optimal bandwidth required for x number of video calls in Skype instantaneously by using the QoE metric. Quality of experience is the metric to measure end-user expectations [31].

By the QoE, we can measure the video quality using peak signal-to-noise ratio (PSNR), mean opinion score (MOS), playback smoothness, and video quality measurement (VQM), these metrics come under objective and subjective metrics of QoE. So, QoE mainly focuses on making the user convenient with the streaming video that means it is related to user satisfactory point of view, where QoS related to making use of network resources efficiently for providing quality service to the user. But the fact is that there is no particular quality of service (QoS) parameter to estimate the user's streaming experience in case of real-time data streaming [32]. Because of this only, to consider the user opinion in video quality evaluation, QoE is evolved. QoE as well as QoS both are not mutually disjointed, but QoE is a superset of QoS. The relation of QoS and QoE is shown in Fig. 45.2. The QoE is influenced by the factors like system level (bandwidth variation, packet loss, delay, jitter, end-user

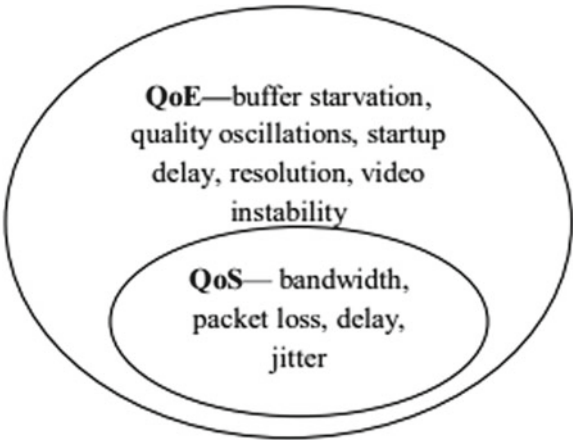


Fig. 45.2 Adaptive streaming process

device configuration, browser), context level (user location, streaming purpose like education, gaming), user level (expectations), and content level (bitrate, resolution, video quality) [33].

45.5 Machine Learning Techniques

Nowadays, any work is coming under automation by the computer. It is possible for the system by using machine learning (ML) algorithms, and these algorithms make the system to take decisions through inference, create and utilize the knowledge, and improve the performance. It predicts and provides recommendations based on the results obtained by processing too large and too complex data sets [28, 34]. ML predicts an optimal solution by analyzing the past data in a short period. ML techniques help solve automation challenges in network setup, control, and management [7–9, 12].

Because of this reason, we use machine learning technology to automatically implement network functions by interacting with internal and external environments. We can select a suitable ML algorithm based on the collected data nature [7, 8]. Generally, ML is classified into three categories as supervised, unsupervised, and reinforcement learning. In supervised learning (SL), the data is already trained that means input is mapped with output. SL is more helpful for the problems of regression and classification. In unsupervised learning (USL), the data is not trained earlier that means we can give the input values for the collected data but not give the output values. USL is most useful for analysis of the clusters. Reinforcement learning (RL) works on the principle of dynamic decision as well as iterative learning. It does not reveal to the learner what action to take but must use continuous trails to determine the actions that produce the output closest to the target.

The operation and management of ML based technologies involve the task of effectively utilizing resources while always best meeting the needs of services and users [1]. They need to understand changes in system status, understand uncertainty, (re)configure networks, anticipate pressing challenges, and propose appropriate solutions promptly. It can also consider nod computing power, bandwidth availability, available energy for resource allocation, as well as management. Therefore, when applied to the network function of the current real-time streaming data application, these classic ML methods can be modified to improve accuracy, reduce complexity, and reduce latency and QoS [12, 35]. [36] proposes a model using a machine learning algorithm to establish the relationship between QoS–QoE. In [12] authors expressing the SL or semi-SL model is more suitable than USL for QoS and QoE services.

45.6 State of the Art on Video Streaming Adaptation Algorithms in QoS and QoE

A lot of work is being done on QoS and QoE in real-time streaming data using adaptive streaming technologies. We discuss a few of them as given in Table 45.1 above. According to recent measurement studies, the research aims to improve real-time streaming data because it is becoming the major traffic of the network [30, 31]. The main reason to increase video traffic over the network is the availability of smart devices in the hands of almost everyone. Although the top layers of the Internet have undergone major changes and are now used as platforms for large-scale video transmission, most Internet traffic is still mostly transmitted over TCP. TCP uses the congestions algorithms which are loss and more latency, suitable for video-on-demand (VOD) streaming not for live that means video conferencing applications require more throughput as well as less time to establish the connection between client and server for the completion of streaming to achieve QoE of the user [37–39].

This research effort aims to exploit the potential of ML technology to design new methods of real-time data flow to overcome the limitations of TCP and other resource constraints. Using features such as bandwidth, processing power, and traffic latency, the ML design decision algorithm is used to dynamically adapt the recommendations to manage congestion and enhance smooth data flow. A request management framework for heterogeneous networks is to improve the quality of experience (QoE) of users. This proposes a QoE development method based on ML technology to perform load balancing for the global request management model [7, 32, 42]. It estimates or predicts the user's experience and expected quality range by collecting sets of data from heterogeneous networks.

45.7 Conclusion

Real-time transmission of media data over the network needs more bandwidth for high throughput to provide lost data instantly without interruption to the end user. Due to strict delay and jitter requirements, more and more multimedia applications in today's computer networks are unable to use existing technologies based on TCP congestion control. As the level of congestion increases, the quality of these multimedia applications declines to the point where users cannot receive content. Thus, a new generation of Internet-based multimedia applications using machine learning technology has entered the network design and management to pursue the new potential of network systems.

Table 45.1 Comprehensive information on HAS and DASH on QoS and QoE

References	Objective	Approach	Result
[19]	Streaming traffic monitoring for video-on-demand services (HAS)	Stochastic model	The model is best suited and performed data analysis for smooth streaming without congestion
[18]	Overcome the hurdles that existed in HAS in case of bandwidth fluctuations in a short period	A novel bitrate control technique	HAS becoming more powerful and perfect in a network where fluctuations happening
[31]	Make the user happy by providing expected quality video, even though network conditions are fluctuating	Developed an algorithm which takes the maximum bit rate video that does not cross the current bandwidth	Meets the user's expected QoE
[37]	Obtain maximum throughput for delay-sensitive applications communication	Estimation method using protocols of closed-loop flow control	Provide better QoS and QoE to the user
[16]	Real-time identification algorithm over networks for streaming of video with dynamic changes occurs in the network	Bandwidth vary identification method, a Bayesian approach	It identifies congestion in advance and makes video streaming seamlessly to the user's expected level
[11]	Implement dynamic HTTP streaming to adapt to the transport protocol's congestion window size, in the infeasibility of TCP's congestion and error control for low latency, high-rate communications services leads to significant underutilization of the available bandwidth	Novel transport protocol—Predictably reliable teal-time transport (PRRT), a protocol layer that efficiently supports the reliability required by multimedia services under their specific time constraint	It achieves the optimal trade-off between those mechanisms under strict delay constraints to optimize bandwidth utilization. It evaluates PRRT's bandwidth utilization and compares it to recent media streaming standards such as HTTP-DASH (dynamic adaptive streaming over HTTP)
[40]	Minimizing congestion and delay at both side of communication	Cost-based distributed technique for controlling congestion	Provides good efficiency in congestion and delay
[41]	Handling congestion and delay queuing	A novel technique for controlling congestion	Yields good results and hence employed in Google chrome also

(continued)

Table 45.1 (continued)

References	Objective	Approach	Result
[42]	Achieve control over congestion in public networks	Fairness-driven queue management algorithms	The survey gives a comprehension of information regarding challenges in congestion control in a network of heterogeneous types

References

1. Martin, A., Egaña, J., Flórez, J., Montalbán, J., Olaizola, I.G., Quartulli, M., Viola, R., Zorrilla, M.: Network resource allocation system for QoE-aware delivery of media services in 5G networks. IEEE Trans. Broadcast. **64**(2) (2018)
2. Cisco Systems, Inc.: Cisco Visual Networking Index: Forecast and Methodology, 2016–2021, White Paper, 2017. (Online). Available: https://www.reinvention.be/webhdfs/v1/docs/complete-white-paper-c11-481360.pdf (2017)
3. Jacobson, V., Frederick, R., Casner, S., Schulzrinne, H.: Real-TimeTransport Protocol (RTP). https://www.ietf.org/rfc/rfc3550.txt (2017)
4. Adobe: Real-Time Messaging Protocol(RTMP) (2017)
5. Schulzrinne, H.: Real Time Streaming Protocol version 2.0. https://tools.ietf.org/html/rfc7826 (2017)
6. Goldberg, J., Westerlund, M., Zeng, T.: A Network Address Translator (NAT) Traversal Mechanism for Media Controlled by the Real- Time Streaming Protocol (RTSP) (2017)
7. Ibarrola, E., Davis, M., Voisin, C., Close, C., Cristobo, L.: A machine learning management model for QoE enhancement in next-generation wireless ecosystems. In: IEEE ITU Kaleidoscope: Machine Learning for a 5G Future (ITU-K) (2018)
8. Kafle, V.P., Fukushima, Y., Martinez-Julia, P., Miyazawan, T.: Cosideration on automation of 5G network slicing with machine learning. In: IEEE ITU Kaleidoscope: Machine Learning for a 5G Future (ITU-K) (2018)
9. Abadi, M. et al.: TensorFlow: large-scale machine learning on heterogeneous distributed systems. (Online). Available: https://arxiv.org/abs/1603.04467 (2016)
10. Ferreira, L.E.B., Gomes, H.M., Bifet, A., Oliveira, L.S.: Adaptive random forests with resampling for imbalanced data streams. IEEE International Joint Conference on Neural Networks (IJCNN) (2019)
11. Gorius, M., Shuai, Y., Herfet, T.: Dynamic media streaming under predictable reliability. In: IEEE international Symposium on Broadband Multimedia Systems and Broadcasting (2012)
12. Aroussi, S., Mellouk, A.: Survey on machine learning-based QoE-QoS correlation models. In: International Conference on Computing, Management and Telecommunications (ComManTel), pp. 200–204 (2014)
13. Ibarrola, E., Saiz, E., Zabala, L., Cristobo, L., Xiao, J.: A new global quality of service model: QoXphere. IEEE Commun. Mag. **52**, 193–199 (2014)
14. Martín, V., Cabrera, J., García, N.: Design, optimization and evaluation of a Q-learning HTTP adaptive streaming client. IEEE Trans. Consum. Electron. **62**(4), 380–388 (2016)
15. Gadaleta, M., Chiariotti, F., Rossi, M., Zanella, A.: D-DASH: a deep Q-learning framework for DASH video streaming. IEEE Trans. Cognitive Comm. Netw. **3**(4),703–718 (2017)
16. Javadtalab, A., Semsarzadeh, M., Khanchi, A., Shirmohammadi, S., Yassine, A.: Continuous one-way detection of available bandwidth changes for video streaming over best-effort networks. IEEE Trans. Instrum. Meas. **64**(1) (2015)
17. Gomez-Uribe, A., Hunt, N.: The Netflix recommender system: algorithms, business value, and innovation. ACM Trans. Manage. Inf. Syst. **6**(2016)

18. Tian, G., Liu, Y.: Towards Agile and smooth video adaptation in HTTP adaptive streaming. IEEE/ACM Trans. Netw. **24**(4), 2386–2399 (2016)
19. Waldmann, S., Miller, K., Wolisz, A.: Traffic model for HTTP-based adaptive streaming. In: IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) (2016) (2017)
20. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol. **13**(7), 560–576 (2003)
21. Schwarz, H., Marpe, D., Wiegand, T.: Overview of the scalable video coding extension of the H.264/AVC standard. IEEE Trans. Circuits Syst. Video Technol. **17**(9), 1103–1120 (2007)
22. Sullivan, G.J., Ohm, J.-R., Han, W.-J., Wiegand, T.: Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol. **22**(12), 1649–1668 (2012)
23. Boyce, J.M., Ye, Y., Chen, J., Ramasubramonian, A.K.: Overview of SHVC: scalable extensions of the high efficiency video coding standard. IEEE Trans. Circuits Syst. Video Technol. **26**(1), 20–34 (2016)
24. Li, C., Wei, F., Dong, W., Wang, X., Liu, Q., Zhang, X.: Dynamic structure embedded online multiple-output regression for streaming data. IEEE Trans. Pattern Anal. Mach. Intell. **41**(2) (2019)
25. Yusuf-Asaju, A.W., Dahalin, Z.M., Ta'a, A.: Framework for modelling mobile network quality of experience through big data analytics approach. J. Inf. Commun. Technol. (JICT) **17**, 79–113 (2018)
26. Tripathi, A., Ashwin, T.S., Guddeti, R.M.R.: EmoWare: a context-aware framework for personalized video recommendation using affective video sequences. IEEE Access **7** (2019)
27. Bogale, T.E., Wang, X., Le, L.: Machine intelligence techniques for next-generation context-aware wireless networks. ITU J. ICT Discoveries (Special Issue No. 1) (2018)
28. Jiang, C. et al.: Machine learning paradigms for next-generation wireless networks. IEEE Wirel. Commun. **24**, 98–105 (2017)
29. Kato, N., et al.: The deep learning vision for heterogeneous network traffic control: proposal, challenges and future perspective. IEEE Wirel. Commun. **24**(3), 146–153 (2017)
30. Xu, Y., Yu, C., Li, J., Liu, Y.: Video telephony for end-consumers: measurement study of Google+, iChat, and Skype. IEEE/ACM Trans. Netw. **22**(3), 826–839 (2014)
31. Qadir, Q.M., Kist, A.A., Zhang, Z.: A Novel traffic rate measurement algorithm for quality of experience-aware video admission control. IEEE Trans. Multimedia **17**(5) (2015)
32. Vega, M.T., Perra, C., Turck, F.D., Liotta, A.: A review of predictive quality of experience management in video streaming services. IEEE Trans. Broadcast. **64**(2) (2018)
33. Hoßfeld, T., Schatz, R., Biersack, E., Plissonneau, L.: Internet video delivery in YouTube: from traffic measurements to quality of experience (2013)
34. De Grazia, M.D.F., Zucchetto, D., Testolin, A., Zanella, A., Zorzi, M., Zorzi, M.: QoE multi-stage machine learning for dynamic video streaming. IEEE Trans. Cognitive Comm. Netw. **4**(1), 146–161 (2018)
35. Elwerghemmi, R., Heni, M., Ksantini, R., Bouallegue, R.: Online QoE prediction model based on stacked multiclass incremental support vector machine. In: IEEE 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO) (2019)
36. Focus Group on Machine Learning for Future Networks including 5G. https://www.itu.int/en/ITUT/focusgroups/ml5g/Pages/default.aspx (2018).
37. Lübben, R., Fidler, M.: Service curve estimation-based characterization and evaluation of closed-loop flow control. IEEE Trans. Netw. Serv. Manage. **14**(1) (2017)
38. Kanrar, S., Mandal, N.K.: Video traffic flow analysis in distributed system during interactive session. Adv. Multimedia **2016**, 14. Article ID 7829570 (2016)
39. Balaouras, P., Stavrakakis, I.: A self-adjusting rate adaptation scheme with good fairness and smoothness properties. Comput. Netw. **48**, 829–855 (2005)
40. D'Aronco, S., Toni, L., Mena, S., Zhu, X., Frossard, P.: Improved utility-based congestion control for delay-constrained communication. IEEE/ACM Trans. Netw. **25**(1) (2017)

41. Carlucci, G., Cicco, L.D., Holmer, S., Mascolo, S.: Congestion control for web real-time communication. IEEE/ACM Trans. Netw. **25**(5) (2017)
42. Abbas, G., Halim, Z., Abbas, Z.H.: Fairness-driven queue management: a survey and taxonomy. IEEE Commun. Surv. Tutorials **18**(1) (2016)

Chapter 46
A Study on the State of Artificial Intelligence in the Design Phase of Software Engineering



Mriganka Shekhar Sarmah, Jediael Meshua Sumer, Marlom Bey, and Bobby Sharma

Abstract The areas of software engineering and artificial intelligence are among the most sought-after fields in the computer science community. This paper assesses the current scenario of implementing artificial intelligence in the design phase of software engineering and attempts to offer a viable explanation to support the conclusion of the assessment.

46.1 Introduction

In this day and age, traditional software design has its limits. In traditional SE, the approach towards building a software is manual in nature, which makes it prone to human errors such as oversight and is also inefficient when complicated projects are considered [1, 2].

Another reason for errors may be the ambiguity of the textual requirements which may not be exactly translated according to the perspective of the customer [3].

Therefore, in order to minimize human error and increase the efficiency of software design and to deal with ambiguity of textual requirements of the customer, AI techniques can be implemented, mainly in the stages of design, testing and code generation [4, 5].

Hence, the prospects of automation/AI in the phase of software design must be discussed, and related problems must be dealt with, which are the reliability and accuracy of automation and how much control it takes away from a software developer.

This paper compares related works regarding how they either tackle this problem or suggest ways to tackle this problem.

M. S. Sarmah (✉) · J. M. Sumer · M. Bey · B. Sharma
Department of Computer Science and Engineering, Assam Don Bosco University, Guwahati, Assam, India

B. Sharma
e-mail: bobby.sharma@dbuniversity.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_46

The paper is organized into the following sections. Section 46.2 discusses works related to the stated topic, Sect. 46.3 discusses the analysis and results of existing work, and finally with Sect. 46.4, the paper is summarized and concluded.

46.2 Related Works

In [4], the author discusses various commonalities between the fields of AI and SE, current status and upcoming trends. Intersections between AI and SE are discussed, especially in the fields of agent-oriented software engineering, knowledge-based engineering systems, computational intelligence and knowledge discovery of data and ambient intelligence. It also gives a short description about AI and software engineering and the possibilities of interaction with one another through various common points of contact.

In [5], the author discusses the differences between the software engineering process and AI and the hostilities that exist in software engineers towards automation in developing a software. It further discusses the contribution of AI to the field of SE in fields like automatic programming and some projects like REFINE and Programmer's Apprentice (an MIT project) and the problem areas of SE where AI can be implemented like requirements analysis and definition, process modelling, process support and project planning.

The author states that one of the reasons for the existing hostility towards AI is that AI researchers are themselves software engineers, and therefore their method to integrating AI in software engineering follows a typical and traditional approach which precedes the advancements that AI and its endless possibilities can bring to the field of software engineering.

In [6], the author uses the terminology defined in IEEE 12,207 standards for software engineering to describe the development processes of requirements analysis, software architecture design and coding and testing and AI techniques that can be implemented along with them. The paper also talks about open problems in SE that can be solved using AI so as to improve the process of SE in this day and age, where software is getting more modular and complex to develop.

In [7], the author deals with a taxonomy called AI in SE Application Levels or AI-SEAL that classifies 15 papers from previous editions of the RAISE workshop. It considers the context in which AI is being applied, i.e. "when" and "on what" AI is being applied. After classification, the papers also assign levels of automation to the papers.

In [8], the author gives a brief overview of SE and expert systems in artificial intelligence and how expert systems can be used to automate the programming process and code generation via the use of genetic programming.

This paper also shows the absence of risk management in AI-based systems due to the way they work.

In [1], the author discusses the translation of user requirements into design diagrams and describes it as a daunting task for a designer, as that person has to translate textual requirements into a diagrammatic (UML) form. This paper gives a basic overview of existing systems regarding and how they convert the user requirements using NLP into UML diagrams.

They developed their own tool, DC Builder, and presented in the paper a diagrammatic representation and a heuristic set of rules in implementing NLP on user requirements. Finally, they evaluated and compared various tools mentioned and developed, with one another.

In [2], the author describes the extraction of UML diagrams from textual requirements by requirements of engineers as a daunting task and the time and effort spent on this justify a tool to automate this process which brings the author to propose a tool called RAPID which uses NLP in an efficient manner to extract design diagrams from input textual user requirements.

The methodology states various NLP technologies, algorithms and rules to extract class information from the textual requirements.

In [3], the author talks about the ambiguity of textual requirements and the usage of NLP and domain ontology to generate UML diagrams from the said textual requirements.

This paper also talks about existing systems and then approaches the problem using their own method called RAUE to extract UML diagrams from textual requirements.

In [9], the author looks at existing systems which translate textual requirements into UML diagrams. After that the author proposes his own approach called RACE which is an improvement on the TCM system. The methodology of the RACE system is discussed with its algorithms and rules to identify classes, attributes and relationships.

The author concludes the paper stating RACE being an advanced approach towards extraction of UML diagrams from textual requirements using a human-centred UI and what the system did not support.

In [10], the author introduces an automatic test data generator called ATGen. ATGen is based on constraint logic programming and symbolic execution. Developing software includes a number of testing methods. The main technique that is currently used in the industry is dynamic software testing where the software is executed using test data. Dynamic testing can be performed using automatic tools, which are automation of administrative tasks, automation of mechanical tasks and automation of test generation tasks.

ATGen is a prototype testing tool implemented using the ECLiPSe constraint logic programming environment and consists over 5000 lines of commented Prolog code. The current area of application of ATGen is the automatic generation of test data to achieve 100% decision coverage for programmes written in SPARK Ada. In decision testing, the aim is to test all decision outcomes in the programme. The aim is to generate a test data suite achieving 100% decision coverage. However, initial results of ATGen have proved that at the time this research was conducted, and it was impossible to experiment with ATGen using industrial SPARK Ada code as such code

was not usually made available even for research due to its safety critical aspects. Also, the test results show that ATGen shows wide performance variation between successive runs, particularly for programmes with loops. The initial results of ATGen were promising, and the overall efficiency of the algorithm is under improvement.

In [11], the author describes an approach towards extracting UML diagrams from natural language using NLP via a tool called CM Builder. CM Builder is a graphical CASE tool that does surface analysis of text to propose candidates for class, attributes and relationship and domain-independent semantic analysis to automatically extract the candidates and finally represent them via UML diagrams.

The tool also includes capacity to evaluate its candidate classes. The author also states benefits of object-oriented analysis of the tool and its scope of improvement.

In [12], the author explored the reasons for the lack of impact in important areas in which AI has been expected to significantly affect real-world software engineering. The session approached the failures of AI in software engineering, looking at the matter through a common cause reliance on isolationist technology and approaches, rather than upon creating additive technology and approaches that can be integrated with other existing capabilities.

The isolationism has been manifested in several areas, in essence, the market has rejected the isolationist and egocentric approach to implementing AI in software engineering, and that the whole system should be developed and executed in generalized workstations and PCs.

In [13], the author describes recent state-of-the-art approaches to automatically generate regular expressions from natural language specifications. Given that these approaches use only synthetic data in both training datasets and validation/test datasets, a natural question arises: Are these approaches effective to address various real-world situations? To explore this question, in this paper, a characteristic study on comparing two synthetic datasets is used by the recent research, a real-world dataset is collected from the Internet, and an experimental study on applying a state-of-the-art approach on the real-world dataset is conducted. The study results suggest the existence of distinct characteristics between the synthetic datasets and the real-world dataset, and the state-of-the-art approach (based on a model trained from a synthetic dataset) achieves extremely low effectiveness when evaluated on real-world data, much lower than the effectiveness when evaluated on the synthetic dataset.

In [14], the author looks at previous work done in the field of converting natural language requirements into design diagrams and points us the various issues that can be encountered in this process.

The author points out the issues in the context of software engineering, then in the context of natural language and finally proposes a methodology to solve these problems to get good quality UML diagrams.

In [15], the author introduces artificial intelligence to software engineers and conversely, software engineering to artificial intelligence workers. It further highlights the contrast between the two fields and further accentuates their differences in the problems that they attempt to solve, their methodologies and the tools and techniques. The author strongly believes that the work of software engineers and artificial intelligence workers are similar, and the fusion of the two fields is essential for

computer science as a field to move forward; however, the author also acknowledges the ridge that exists between the two communities. The paper does not, however, provide the "*how*" on which AI and software engineering can come together, other than details of the core components of both fields.

The paper acknowledges that to move forward with "user-friendliness", the system needs to allow natural language input and output, and natural language processing is one of the most researched areas in AI. In other words, a system needs AI to further improve its usability.

Conversely, AI is in need of proper models to address its problems. These models are already evident in conventional systems, and it is clear that AI systems need some standardization in order to widen their applications.

Thus, there exists a requirement for both fields to merge on some common ground.

46.3 Results and Analysis of Existing Work

From the study done above, a few things are apparent. Firstly, when design diagrams made by a software engineer are considered, three factors come into play.

- Ambiguity and complexity of NL textual requirements [13].
- Complex nature of the project due to increasing complexity of technology and size of the project.
- Possibility of human error when designing those systems from textual requirements.

The above results in the problem of defining the accuracy of a system, be it designed by a person or by an AI-based tool, as there is no point of references. This makes it imperative to create a dataset to compare any diagram to measure its accuracy.

Another way it can be done is to design an algorithmic model that can analyse input textual requirements to give appropriate output, but the amount of work involved in making it foolproof would make it an insurmountable challenge.

Another aspect of translating user requirements into a design diagram is that when the textual requirements are complicated, extracting context requires the use of domain ontology along with defined semantic rules. When using this, problems will be encountered when very complex sentences are considered.

Secondly, when AI in system design is considered, due to its automatic nature, the problem of a system engineer not having enough freedom over the creation of design diagrams is encountered. Hence, there exists a need to carefully define the level of automation of such tools. This can be done by making the end results of a diagram generator interactive or editable, as can be seen in various smart design diagram tools.

Table 46.1 shows the comparison of the availability of features of various tools that generate design diagrams from user input textual requirements. It can be inferred

Table 46.1 Evaluation of tools' functionality. Adapted and modified from [1]

Support	CM Builder [11]	LIDA	GOOAL	NLOOML	DC Builder [1]	RACE [9]	RAUE [3]	RAPID [2]
Classes	Yes	User	Yes	Yes	Yes	Yes	Yes	Yes
Attributes	Yes	User	Yes	Yes	Yes	Yes	Yes	No
Methods	No	User	Yes	Yes	No	Yes	No	No
Associations	Yes	User	Semi-NL	No	Yes	Yes	Yes	Yes
Multiplicity	Yes	User	No	No	No	No	Yes	No
Aggregation	No	No	No	No	Yes	Yes	Yes	Yes
Generalization	No	No	No	No	Yes	Yes	Yes	Yes
Instances	No	No	No	No	No	No	No	No

from the table that there are very few tools which fully extract relationships from textual requirements.

Another observation made from the tools mentioned is that although they can create class diagram, they cannot deal with instances to create object diagrams.

46.4 Conclusion and Future Work

From the above survey on the papers, it is evident that artificial intelligence has a wide scope of application in the field of software engineering and has the capability to make a software engineer's work easier and faster. However, AI systems have to undergo further improvements to fully exploit their capabilities in the field of software engineering.

The findings of this paper can be further analysed and utilized to design methodologies that make generation of UML design diagrams from input user requirements far more efficient. Furthermore, more research has to be done, aimed towards the use of machine learning to automate the task of diagram generation.

References

1. Herchi, H., Abdessalem, W.B.: From user requirements to UML class diagram (2012). arXiv preprint arXiv:1211.0713
2. More, P., Phalnikar, R.: Generating UML diagrams from natural language specifications. Int. J. Appl. Inf. Syst. Found. Comput. Sci. **1**(8), 19–23 (2012)
3. Joshi, S.D., Deshpande, D.: Textual requirement analysis for UML diagram extraction by using NLP. Int. J. Comput. Appl. **50**(8), 42–46 (2012)
4. Rech, J., Althoff, K. D.: Artificial intelligence and software engineering: status and future trends. KI **18**(3), 5–11 (2004)

5. Sommerville, I.: Artificial intelligence and systems engineering. In: Prospects for Artificial Intelligence: Proceedings of AISB'93, 29 Mar, 2 Apr 1993, Birmingham, UK, 17, 48 (1993)
6. Ammar, H.H., Abdelmoez, W., Hamdi, M.S.: Software engineering using artificial intelligence techniques: Current state and open problems. In: Proceedings of the First Taibah University International Conference on Computing and Information Technology (ICCIT 2012), Al-Madinah Al-Munawwarah, Saudi Arabia, p. 52 (2012)
7. Feldt, R., de Oliveira Neto, F.G., Torkar, R.: Ways of applying artificial intelligence in software engineering. In: 2018 IEEE/ACM 6th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE), pp. 35–41. IEEE (2018)
8. Raza, F.N.: Artificial intelligence techniques in software engineering (AITSE). In: International MultiConference of Engineers and Computer Scientists (IMECS 2009), vol. 1 (2009)
9. Ibrahim, M., Ahmad, R.: Class diagram extraction from textual requirements using natural language processing (NLP) techniques. In: 2010 2nd International Conference on Computer Research and Development, pp. 200–204. IEEE (2010)
10. Meudec, C.: ATGen: automatic test data generation using constraint logic programming and symbolic execution. Softw. Test. Verification Reliab. **11**(2), 81–96 (2001)
11. Harmain, H.M., Gaizauskas, R.: CM-Builder: an automated NL-based CASE tool. In: Proceedings ASE 2000. 15th IEEE International Conference on Automated Software Engineering, pp. 45–53. IEEE (2000)
12. Balzer, R., Fikes, R., Fox, M., McDermott, J., Soloway, E.: AI and software engineering: will the twain ever meet?. In: AAAI, pp. 1123–1125 (1990)
13. Zhong, Z., Guo, J., Yang, W., Xie, T., Lou, J. G., Liu, T., Zhang, D.: Generating regular expressions from natural language specifications: are we there yet? In: Workshops at the 32nd AAAI Conference on Artificial Intelligence (2018)
14. Sharma, N., Yalla, D.P.: Issues in developing UML diagrams from natural language text. In: Proceedings of WSEAS Conference, Recent Advances in Telecommunications, Informatics and Educational Technologies, pp. 139–145
15. Ford, L.: Artificial intelligence and software engineering: a tutorial introduction to their relationship. Artif. Intell. Rev. **1**(4), 255–273 (1987)

Chapter 47
S2S Translator by Natural Language Processing



Srikanth Bethu, Suresh Mamidisetti, S. Bhargavi Latha, and B. Sankara Babu

Abstract Correspondence by means of motions is a visual language that is used by in need of a hearing-aided people as their local language. Not in the least like acoustically passed on sound models, signal-based correspondence uses non-verbal correspondence and manual correspondence to easily pass on the contemplations of a person. It is practiced by simultaneously uniting hand shapes, heading and advancement of the hands, arms or body, and outward appearances. In this paper, we will endeavor to develop a programmed model which changes over voice to motion-based correspondence and besides correspondence by means of motions to voice/text. We may be using talk-to-message programming interface (Python modules or Google programming interface) and a while later using the semantics of natural language processing to breakdown the substance into more diminutive sensible pieces which requires AI as a segment. Instructive assortments of predefined motion-based correspondence are used as the information with the objective that the item can use manmade mental ability to show the changed over sound into the signal-based correspondence.

47.1 Introduction

Motion-based correspondence is a language absolutely autonomous and unquestionable from English. It contains all the basic features of language, with its own rules for enunciation, word game plan, and word demand. While every language has techniques for hailing different limits, for instance, representing a request rather than saying something, vernaculars differentiate in how this is done. For example, English speakers may represent a request by raising the pitch of their voices and by adjusting word demand; sign language [1] customers represent a request by creating

S. Bethu (✉) · S. B. Latha · B. S. Babu
Department of Computer Science and Engineering, GRIET, Hyderabad, Telangana 500090, India

S. Mamidisetti
Department of Technical Education, Government Polytechnic College, Hyderabad 500028, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_47

a ruckus, expanding their eyes, and slanting their bodies forward. Correspondingly moreover with various lingos, unequivocal techniques for conveying considerations in signal-based correspondence move as much as customers themselves. Despite solitary differences in enunciation, correspondence by means of motions has commonplace accents and tongues; comparably as certain English words are verbally communicated particularly in different bits of the country, motion-based correspondence has regional assortments in the musicality of checking, address, slang, and signs used.

Other sociological components, including age and sexual direction, can impact correspondence through signals usage and add to its combination, correspondingly moreover with conveyed in vernaculars. Finger spelling is a bit of correspondence through marking and is used to light up English words. In the finger spelled letters all together, each letter identifies with a specific hand shape. Finger spelling is as often as possible used for real names or to show the English word for something.

Correspondence through marking is something which is not fathomed by various people. Generally scarcely any people have the prolog to it which makes it hard for people who talk with it to have the alternative to give their expects to others. Disregarding the way that some may appreciate bits of it in any case, they cannot grasp its whole setting. Considering the circumstance as we have right now, nearly deaf people are commonly prevented from securing standard correspondence with others, they have to rely upon an interpreter or some visual correspondence, similarly by virtue of ordinary people as well.

47.2 Literature Review

Quoc et al. [2] in 2018 proposed a work on highlight talk understanding. Emphasis is used to perceive connected with and unfocused bit of a verbalization, and it is useful in misheard conditions in which speakers must repetitive the most huge words or articulations. The proposed approach to manage handle steady emphasis levels which relies upon course of action models and besides they have joined the machine and complement understanding into the single model. They have proposed hard thought emphasis talk understanding and joint model. In the hard thought complement translation, it can unravel reliable highlight loads without quantization. Furthermore, in the joint model, it unravels the translation and together interprets words and complement with single word delay. They have used LSTM-based encoder-decoder to handle the issues in single model of hard thought progression-to-plan model.

Taylor et al. [3] in 2018 proposed a work on talk affirmation by using disguised Markov model. They consider the translation language from English to Indonesian. The request procedure used is the Mel repeat cepstral coefficients (MFCC) and hidden Markov model (HMM). The proposed structure changes over talk to text and usages the current Google understanding or Microsoft translation for translation. The talk sign will be dealt with by using MFCC. The count used here is k-means estimation.

Poornakala and Maheshwari [4] in 2016 proposed a work talk translation from English to Tamil language [5]. Talk affirmation structures seeing English talk by

methods for a voice affirmation device. English talk make an understanding of the substance into English using the conversation mix, in the wake of changing over it to message, independent, and the words put to the side in the informational collection in case it energizes the Tamil substance put to the side in the data set. The English substance is changed over to Tamil substance which is made by the machine translation structure, and a while later, the substance will be appeared on the screen. It is expected to decode up to 60 words in the Android application. The estimation used here is HMM count.

Shin et al. [6] in 2015 proposed a work on talk-to-talk translation humanoid robot. The proposed structure is helpful for English-talking patients; they can uncover their issues to Korean authorities. It can deal with the issues of various patients who do not have the foggiest thought regarding the Korean language despite the way that they can portray their issues to pro which override the activity of human workers. It uses CMU Sphinx-4 gadget for talk affirmation. English–Korean understanding relies upon rule-based translation.

47.3 Methodology and Implementation

Image processing method infers treatment of a propelled picture by techniques for a PC. It can in like manner be described as a technique for using PC computations, in order to get improved picture either to remove some supportive information or for some other kind of task.

Speech Recognition: Talk affirmation is a subfield of programming building and historical underpinnings that make methods and advancements that enable the affirmation and translation [7] of conveyed in language into text by PCs. It is in any case called automatic talk affirmation (ASR), PC talk affirmation, or talk to message (STT). Talk affirmation systems require "getting ready" where an individual speaker gets text or fundamental language into the structure. The structure assesses the individual's unequivocal voice and uses it to change the affirmation of that individuals talk, realizing extended exactness. Systems that do not use getting ready are called speaker-free structures. Talk affirmation applications fuse various things, for instance, voice dialing, call coordinating, search catchphrases, clear data area, preparation of sorted out reports, choosing speaker characteristics, and talk to message getting ready.

Models, Methods, and Algorithms: Both language showing and acoustic exhibiting are noteworthy bits of present-day quantifiably based talk affirmation computations. Covered Markov models (HMMs) are comprehensively used in various structures. Language showing is moreover used in various other trademark language taking care of utilizations, for instance, quantifiable machine understanding or document gathering. Markov models: Present-day talk affirmation systems rely upon hidden Markov models [8]. These are authentic models that yield a progression of pictures or sums. HMM is used in talk affirmation because a talk sign can be viewed as a piecewise fixed sign or a short period of time fixed sign.

From Fig. 47.1, the modules' frameworks are created for speech and text to convert into sign language using sign language translator [9]. In Fig. 47.1, the image is taken as input into model, and output generated is voice output using natural language conversion process. From Fig. 47.2, it showed the connectivity of preprocessed input image to preprocessed signed image. In Fig. 47.2, the voice is taken as input, and sign language image is generated output.

Getting the Dataset: Datasets of predefined sign language were taken from Kaggle. It had approximately 10,000 images corresponding to each alphabet including both

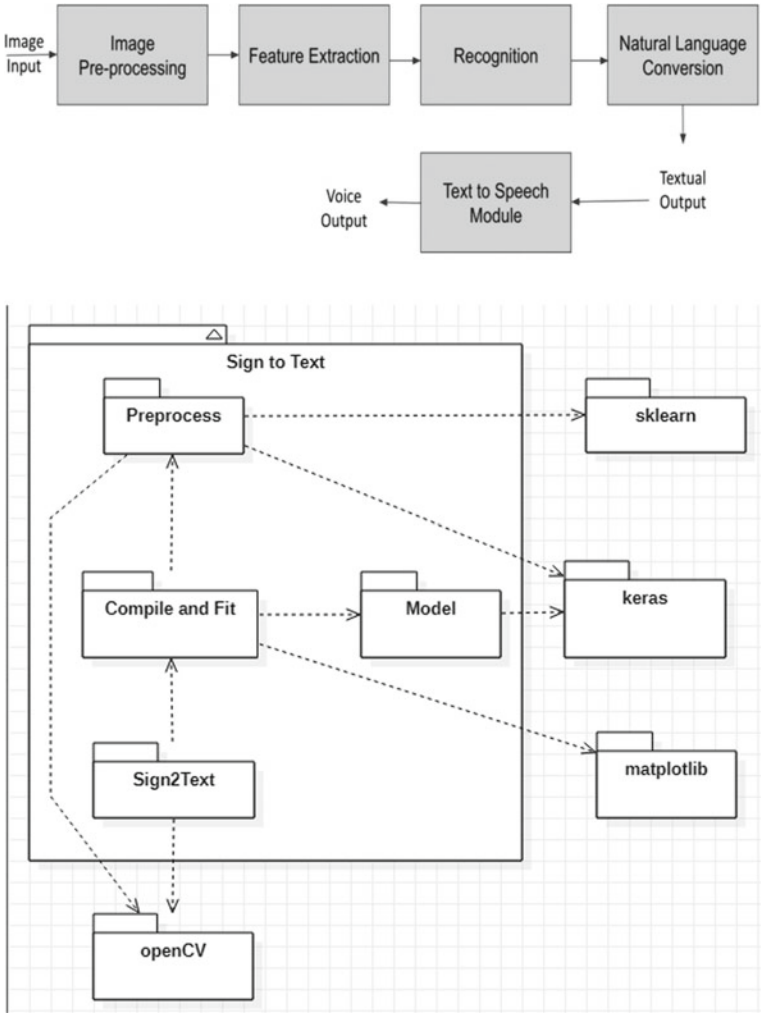


Fig. 47.1 Text-to-speech module creation and the sign-to-text process is created using OpenCV by Matplotlib

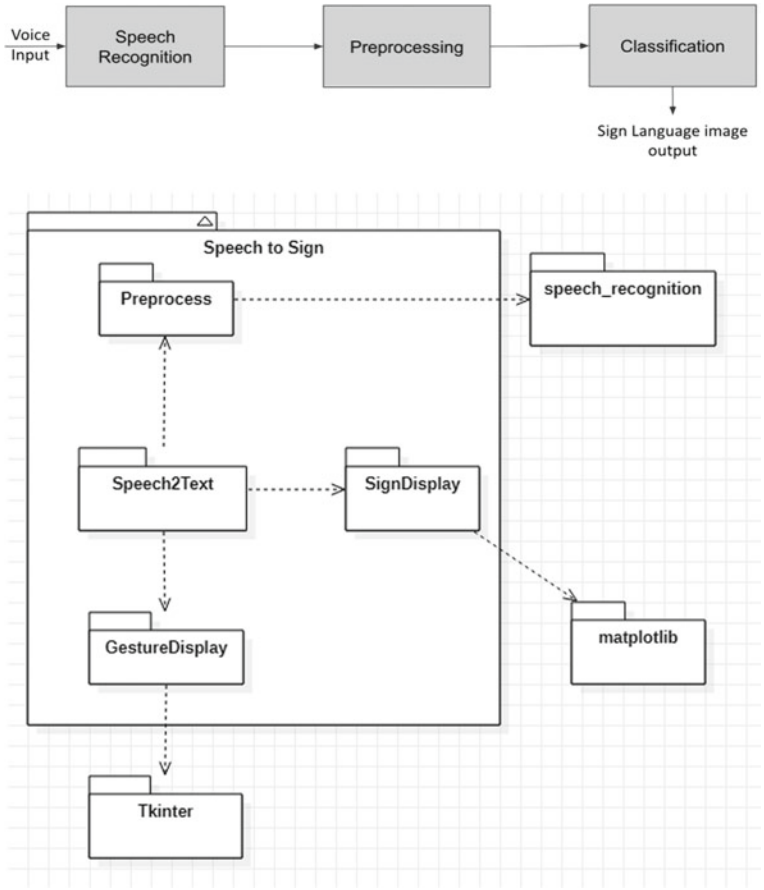


Fig. 47.2 Sign language output module creation and the speech-to-sign process is created using OpenCV by Matplotlib

training and test datasets. An image corresponding to each alphabet was taken and displayed using Matplotlib, so that users can see what the corresponding sign to each alphabet was. Datasets of some frequent sign language [10] gestures were also gathered and were used similarly as mentioned above.

Pre-processing the Dataset: The training data of the sign images was taken and was cleaned to make it ready for classification. Each image was converted to 64X64, and the color was changed to monochrome.

The image array was then converted to a NumPy array. The array was normalized by converting the values to 'float32′ and dividing it with the total sum of the pixel values, i.e., 255. The labels corresponding to each image was stored in an array as well and was converted to categorical values using the internal function present in keras module. The whole image data was split into training and testing using the

internal function present in scikit-learn module. The shape of the train and test data was printed and returned back for the model to consume.

Creating CNN Model: After preprocessing of the training data, we will create a CNN model which is used for classification. Keras module is used extensively here to create the model. We created a sequential model with 16 nodes in the first layer having the kernel_shape [3,3], and the input shape was [64,64,3]. Activation function used was "Relu". Next layer had 32 nodes having the kernel_shape [3,3] with the similar activation function. We added a MaxPool layer of two dimensions before getting the output to the next layers. The next layer connected to 32 node which was further connected to 64 nodes having a MaxPool2D layer. The output of these 64 nodes was fed to the next layer which had 128 nodes connected further to 256 nodes. After the MaxPool2D layer, batch normalization was done here. We flattened the output of the layers here and set the dropout to "0.5". The last layer of processing had 512 nodes where also kernel_regulizer was set to "0.001". At last, we had the output layer of 29 nodes since we have 29 labels or classes. SoftMax function was used to get the output in terms of integers for easy processing and accurate classification. The model was compiled with "Adam" optimizer, and loss and accuracy metrics were recorded while fitting the data into the model. Summary of the model was printed in case anyone in the future will be able to use it, and the model was saved for further use.

Training the CNN Model: After the model is created, we need to fit our train data into it, so that the model will learn about the data through training, and then, it would be used for further classifying real-time data. We use the "fit" method to give our training data. The batch_size is set as 64 which means that images will be fed to the CNN model in batches with 64 each in one batch. The epoch is set to be 5, i.e., for every five minutes, the data will be fed to the model. The validation split is set as "0.5" which means that the dataset is divided into two parts, one for training and other for validation. After the training is done, an object is returned which has all the metrics that we had put while creating the model. Matplotlib is used to plot the accuracy graphs, loss graphs, and printing out the evaluation accuracy and evaluation loss.

Implementing Speech Recognition: The above procedures were done for converting sign language to speech. Now, coming back to speech-to-sign language conversion, we first need to implement the speech recognition to allow speech to be recognized. We will be using the speech_recognition module for that. Once we receive the speech through microphone, it will be converted to text. After converting to text, we will call the gesture display component or the sign display component after performing some validation checks.

Gesture Display and Sign Display: After receiving the text using speech recognition module, we will call the gesture display component if the text matches the gestures that we have in our dataset. If the gesture matches, then a GIF image will be displayed to the user, indicating the gesture corresponding to the text which was spoken. If the text does not match with any of the gestures, then we will display a series of signs

from the test data of the sign language dataset. The signs will be repeated twice having regular intervals, so that user will be able to perceive them properly.

Implementing Live Classification: After creating the model for converting sign language to speech, we will use it to convert the signs provided through live Web cam feed. We will be using OpenCV to get the images through the live feed. After receiving the images, we will preprocess them similarly as mentioned in Sect. 3. After preprocessing, the image is given to the model to classify. The model will give us a numeric output corresponding to the alphabet that the sign portrays. That alphabet is taken and printed over the screen for user accessibility. After the word is at least of 7 letters long, we will invoke the text-to-speech module to convert that text to speech.

47.4 Results and Discussion

Figure 47.3 shows the dataset uploading into the system model created by neural network layers. Maxpooling and batch normalization are used to filter the data. The total parameters are loaded into the model which are 942.557, trainable parameters which are 942.045, and non-trainable parameters which are 512. Non-trainable parameters are classified by flatten and dense methods.

Figure 47.4 shows the model fitting outputs and validation of its accuracy using the graph. The total number of trained samples is used in the models 74,382, and output generated values are 8265. The evaluation accuracy occurred is 99.75%, and evaluation loss occurred is 0.080104.

Figure 47.5 shows the generation of outputs using OpenCV. The images are the predicted images that are to be to be identified in the final prediction results. Figure 47.6 shows the final prediction of all results in S2S language.

47.5 Conclusion

The concept entitled “S2S Translator” has presented a model for translating sign language to speech and speech to sign language. The model that was developed in this paper has a fair enough accuracy and can be used for further improvisation. For better live classification, an object detection module could have been developed or an already existing one could have been used which will lead to better results in translating. Some kind of support for translating more complex sentences can be further investigated for the future of this paper. The voice output feature which is the output for the sign-to-speech [11] conversion can be further improved.




```
MODEL CREATED
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 64, 64, 16)        448
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 64, 64, 32)        4640
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 21, 21, 32)        0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 21, 21, 32)        9248
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 21, 21, 64)        18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 7, 7, 64)          0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 7, 7, 128)         73856
_________________________________________________________________
conv2d_6 (Conv2D)            (None, 7, 7, 256)         295168
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 2, 2, 256)         0
_________________________________________________________________
batch_normalization_1 (Batch (None, 2, 2, 256)         1024
_________________________________________________________________
flatten_1 (Flatten)          (None, 1024)              0
_________________________________________________________________
dropout_1 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_1 (Dense)              (None, 512)               524800
_________________________________________________________________
dense_2 (Dense)              (None, 29)                14877
=================================================================
Total params: 942,557
Trainable params: 942,045
Non-trainable params: 512
```

Fig. 47.3 Uploading sample dataset and creating a model using CNN layers

```
Train on 74385 samples, validate on 8265 samples
Epoch 1/5
74385/74385 [==============================] - 19s 259us/step - loss: 0.7998 - acc: 0.8382 - val_loss: 0.2221 - val_acc: 0.9843
Epoch 2/5
74385/74385 [==============================] - 15s 196us/step - loss: 0.2047 - acc: 0.9760 - val_loss: 0.1819 - val_acc: 0.9750
Epoch 3/5
74385/74385 [==============================] - 15s 196us/step - loss: 0.1593 - acc: 0.9826 - val_loss: 0.2469 - val_acc: 0.9593
Epoch 4/5
74385/74385 [==============================] - 15s 199us/step - loss: 0.1450 - acc: 0.9860 - val_loss: 0.1758 - val_acc: 0.9811
Epoch 5/5
74385/74385 [==============================] - 15s 203us/step - loss: 0.1449 - acc: 0.9878 - val_loss: 0.0780 - val_acc: 0.9979
```

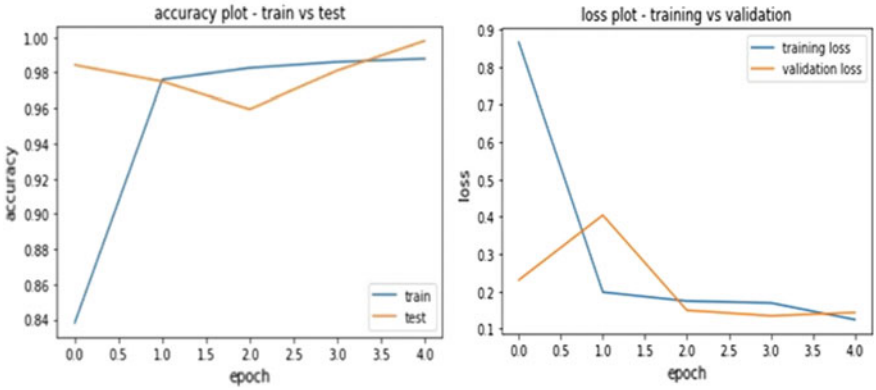


Fig. 47.4 Fitting model output and validating the accuracy plot and loss plot

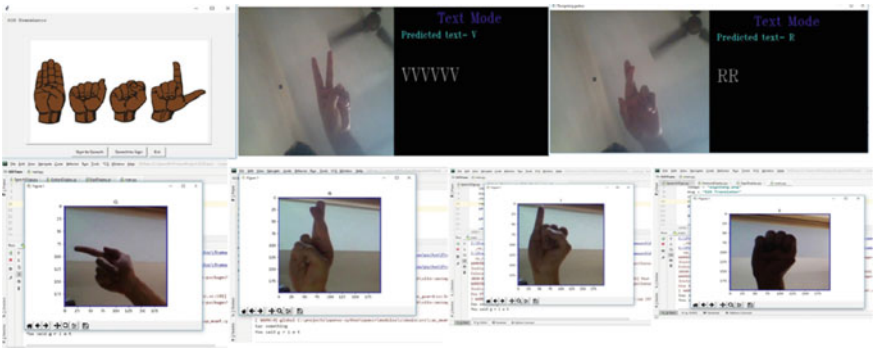


Fig. 47.5 Sign-to-speech and speech-to-text generation output by OpenCV



Fig. 47.6 Final prediction results

References

1. Yang, Q.: Chinese sign language recognition based on video sequence appearance modeling, IEEE. pp. 1537–1542 (2010)
2. Nakamura, S., Sumita, E., Shimizu, T., Sakti, S., Sakai, S., Zhang, J., Finch, A., Kimura, N., Ashikari, Y.: A-STAR: Asia speech translation consortium. In: Proceedings ASJ Autumn Meeting, Yamanashi, Japan (2007), pp. 45–46
3. Taylor, P., Black, A.W., Caley, R.: The architecture of the festival speech synthesis. Proc. Third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis, Blue Mountains, NSW, Australia, 26–29 Nov 1998, pp. 147–151
4. Wagner, R.A., Fischer, M.J.: The string-to-string correction problem. J. Assoc. Comput. Mach. **21**(1), 168–173 (1974)
5. Poornakala, J., Maheshwari, A.: Automatic speech-speech translation form of English language and translate into Tamil language. Int. J. Innov. Res. Sci. **5** (e3) (2016) https://doi.org/10.15680/IJIRSET.2016.0503110.3926
6. Sakti, S., Kimura, N., Paul, M., Hori, C., Sumita, E., Nakamura, S., Park, J., Wutiwiwatchai, C., Xu, B., Riza, H., Arora, K., Li, H.: The Asian network-based speech-to-speech translation

system. ASRU, 2009 and IEEE Explore, 13 Nov 2009, 17 Dec 2009, pp. 507–512 (2009)
7. Alshawi, H., Douglas, S.: Learning dependency transduction models from unannotated examples. Philos. Trans. Roy. Soc. (Ser. A Math. Phys. Eng. Sci.) **358**, 1357–1372 (2000)
8. Moran, D.B., Cheyer, A.J., Julia, L.E., Martin, D.L., Park, S.: Multimodal user interfaces in the open agent architecture. Knowl. Based Syst. **10**(5), 295–303 (1998). Also published in Proceedings of 2nd International Conference on Intelligence. User Interfaces, pp. 61–68 (1997)
9. Hatim, B., Munday, J.: Translation: An advanced Source Book. Routledge, London (2004)
10. Ellis, K., Barca, J.C.: Exploring sensor gloves for teaching children sign language. Adv. Hum. Comput. Interact. 1–8 (2012)
11. Yun, S., Lee, Y.J., Kim, S.H.: Multilingual speech-to-speech translation system for mobile consumer devices. IEEE Trans. Consum. Electron. **60**(3), 508–516 (2014)

Chapter 48
Chemical Plant Liquid Leakage IoT-Based Monitoring



G. P. Hegde, Nagaratna Hegde, and M. Seetha

Abstract The leakage of chemical plant liquid causes hazards to human beings and animals. Liquid would convert into gaseous form and starts spreading everywhere. This paper focuses on automatic monitoring of liquid leakage in chemical plant which causes pollution of environment. This paper presents design pattern of minimum cost liquid leakage monitoring system for real-time detection of quality of liquid passes in longer tubes of chemical plant. This pattern consists of different sensors which are used to measuring physical quality of the liquid. The parameters flow sensor of the liquid can be measured. Storage of large data and its analysis are more useful in various investigations and developments. The Arduino framework is one of the main controller devices in this innovative work, and the sensed data can be configured and stored in cloud server as metadata for the purpose of silent measure of physical variations in liquids.

48.1 Introduction

48.1.1 Liquid Superiority Monitoring

Supply of good quality of liquid is a major problem in industries to control the machinery parts due to variation of liquid properties. Liquid may vary with viscosity, movement of liquid from higher pressure to lower pressure and liquid flow rate. In real-time applications, liquid flow finds in many industrial plants have got major role. Improper injection of liquid to machine also causes the failure of liquid flow in pipes. Hence, it is an essential task to monitor chemical plant leakages of liquid using IoT

G. P. Hegde (✉)
SDM Institute of Technology, Ujire, Mangalore, India

N. Hegde
Vasavi College of Engineering Hyderabad, Hyderabad, India

M. Seetha
G. Narayanamma Institute of Technology and Science, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_48

which is a good solution. In recent days, evolutions in computing and electronics technologies have triggered Internet of Things technology. Internet of Things can be pronounced as the network of electronics devices communicating among them by the help of a controller. IoT is a collection of various communicating devices and transfers the power from one-to-one device. Sensed information can be transmitted from sensors to embedded devices. The controller task distributes the required signal called microchips. It may consume larger power when signals are accessed through appliances. This paper presents a low-cost liquid flow monitoring system, which is a solution for the liquid consumption and quality of liquid. Microcontrollers and sensors are used for that system. Ultrasonic sensor is used to computing liquid level. The other parameters like pH, TDS and viscosity of the liquid can be calculated using different sensors.

This system uses liquid flow sensors which are placed at certain specific distance on liquid pipe. The liquid flow through the pipe can be sensed and sent to Arduino chip and then require output values computed by this board with respect to flow parameters and transferred to clouds via Wi-Fi frame. In this work, analysis of liquid flow has been carried out with respect to instantaneous values, and how much liquid is used in certain time, in a day or in a month, has been noted. Clients access the sensed data which is stored in microchips and delivered to user terminals. The data which is obtained from the sensors can be shown on the Internet and delivers services for broadcasting the data on mobile phones or web application. Figure 48.1 shows typical chemical plant which consists of complex pipe system.



Fig. 48.1 Typical chemical plant [3]

48.2 Earlier Work

48.2.1 Real-Time Work

In the studies from [1], the author proposed that an IoT-based liquid monitoring system measures liquid level in real time. When pipe is connected between two tanks, then liquid level in both the tank goes equal, and then this condition should observe using sensors. This system is based on knowledge of liquid level and can be very important parameter when it comes to the flood occurrences especially in chemical plant pipes. A liquid level sensor is used to detect the desired parameter, and if the liquid level reaches the parameter, the signal will be fed in real time to social network like Twitter. In the [2] that in their frameworks, there were larger amount of works that have been led by the Internet of Things with various kinds of applications as per the specification of end users. The basic idea of the IoT has flooded over in almost all fields of society, and the man power utilization has been made to downstream. The day-by-day IoT information has been updated with basic architecture. Author also presented that using basic architecture, various contexts related problems have been solved and design patterns were developed. All the structural patterns concerned to this architecture involve the different types of works which give applications with respect to Internet of Things.

In the reference [3], writers concluded that continuous monitoring of drinking solution flow in a pipe is analyzed and also presents designing of minimum cost framework quality control liquid monitoring system. This new system measures the chemical disorders in the liquid which enhances the physical state of the liquid. Their work presents a design and progress of a minimum cost framework for real-time monitoring of the liquid quality in Internet of Things (IoT). The system encompasses of many sensors used in measuring physical and chemical constraints of the liquid. In [4], the author has clearly noted that the maintenance person accesses the liquid level from remote place from remote server via remote references. The controlling of motor pump and remote access is a tedious job. The leakage in pipe can be measured through wireless media, and the message is displayed in mobile phones. The sensors used in their system send the signals to watch the level of water in tanks.

48.2.2 Liquid Quality Monitoring

Nikhil Kedia entitled “Liquid Quality Monitoring for Rural Areas-A Sensor Cloud-Based Economical Project”. This paper highlights the entire liquid quality monitoring methods, sensors, embedded design and information dissipation method, role of government, network operator and villagers in ensuring proper information dissipation. It also explores the sensor cloud domain. While automatically enhancing the liquid, quality is not viable at this point, and efficient practice of technology and economic practices can help advance liquid quality and awareness among people

[5]. Jayti Bhatt et al. entitled "Real-Time Liquid Quality Monitoring System". This paper designates to ensure the safe stream of drinking liquid that the quality should be supervised in real time, and for that purpose novel technique Internet of Things (IoT)-based liquid quality monitoring has been planned. In this paper, we present the design of IoT-based liquid quality monitoring system that monitor the quality of liquid in real time. This system consists of some sensors which measure the liquid quality parameter such as pH, turbidity, conductivity, dissolved oxygen and temperature. The measured values from the sensors are processed by microcontroller, and this processed value is conducted remotely to the core controller that is Raspberry Pi using ZigBee protocol. Finally, sensors data can view on Internet browser application using cloud computing [6].

48.3 Existing Framework

48.3.1 Real-Time Work

In industries, liquid may cause several hazards and dangerous unsecured pollution and respiratory problems in society. In real-time work, it requires ongoing modification of liquid resource guiding principle at the levels of international down to individual wells. In most of the places of world, the dirty liquid is used in industries and products are generated. These products are sold out to different marketing places. No proper filling of liquid is used in many areas, and people are not considering the proper awareness. The leakage at certain region in pipe makes loss of liquid and decreases the quantity. So, it will not reach to the destination. So liquid pollution starts increasing throughout the world, and it would cause the problems in various forms. Many people today are suffering from disease due to polluted water intake. This leads to dangerous hazards in the life. It is necessary to maintain the liquid passing mechanism properly in pipes. Reality of various tasks should be carried out to avoid the leakage of liquid.

The liquid flow monitoring system produces thoughtful worldwide problems in real life. Also, normal properties such as volcanoes, algae tints and tremors also change the quality and environmental status of liquid.

48.4 Proposed Framework

48.4.1 System Overview

In this, we present the chemical plant of industries-based lossless liquid flow system and its proper instant monitoring using IoT. The system consists of Arduino device, microcontroller, different type of sensors like liquid flow sensor, pH and viscosity,

sensor and ultrasonic sensor. The Arduino is the microcontroller of the design structure which controls and processes the data created by the sensors. A Wi-Fi module is associated with the Arduino device which helps to transfer the data to the cloud over Internet. The ultrasonic sensor helps to measure the liquid level when the liquid flow reaches certain level, and then the liquid flow can be stopped automatically by turning the motor off or close the liquid flow in pipe by the help of Arduino. The liquid flow sensor measures the quantity of liquid flow through the pipe in a given time; this data will be sent to cloud for storage and analysis purposes. The other sensors like temperature, pH and viscosity sensor measure the liquid quality and help to normalize whether the liquid is useful for equipment operations or any other resolutions. Figure 48.2 shows our proposed framework.

Transformers liquid, like oil, is used to cool the machine. Sometimes, this oil starts to leak, and it can be detected using IoT-based automatic system.

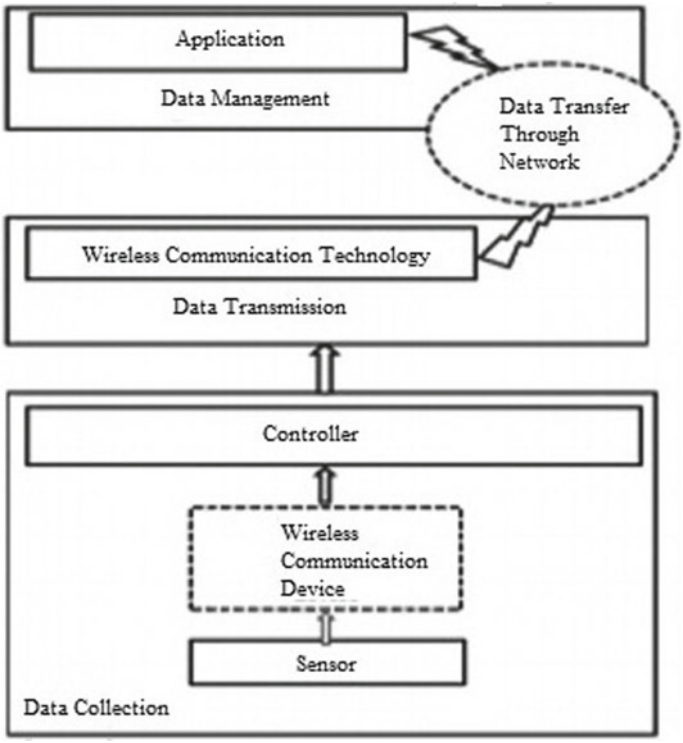


Fig. 48.2 Proposed framework

48.4.2 System Components

Taking about this proposed system, it is clearly shown that it has several components which help to build a liquid monitoring system. The essentail component of the smart home automation system is shown in Fig. 48.2.

Arduino Uno: Arduino is a microcontroller board based on the ATmega328P. It has 14 digital input and output pins, six analog inputs, a 16 MHz quartz crystal, a USB connection, a power jack, an ICSP header and are set button. Figure 48.3 shows typical Arduino board used in this work.

The ESP8266 Wi-Fi Module: It is a self-contained SOC with integrated TCP/IP protocol stack that can give any microcontroller access to your Wi-Fi network. The ESP8266 is capable of either hosting an application or off-loading all Wi-Fi networking functions from another application processor. Each ESP8266 module comes pre-programmed with an AT command set firmware. The ESP8266 module is an extremely cost-effective board with a huge, and ever growing, community. Figure 48.4 shows Wi-Fi module used in this work.

Sensor is used to measure the flow of liquid. This sensor basically consists of a plastic body, a rotor and a sensor. When liquid or water flows through the pipe it influences the pinwheel rotor rotates and wheel speed depends on rate of water flow and is directly proportional to the flow rate of water or liquid.



Fig. 48.3 Arduino Uno board [3]

Fig. 48.4 Wi-Fi module [3]





Fig. 48.5 Liquid flow sensor [3]

Cloud-Based Server: Sensors are placed in equidistant in a longer pipe in industries when the sufficient information and signals are sensed by sensors, and then it has been sent to larger database. They should know the content of cloud server. Based on environmental status, various attributes of liquid like temperature, viscosity and pH values are detected using respective sensors. Various forms of liquid flow can be observed in earlier systems, and it can be measured via fixed sensors. Converting sensed data into positive data computationally increases the response and efficiency of the system. Ultrasonic sensor: Transmitters translate signal into ultrasound, receivers translate ultrasound into electrical signals, and transceivers can both transmit and receive ultrasound. This supports to measure the liquid level.

48.4.3 Internal Overview

In this, we present the theory on real-time monitoring of liquid quality and quantity using IoT. Microcontroller acts as a certain device, and it notices the main signals sent from liquid flow sensor. The programming part of Arduino board is necessarily connected to sensor circuit. The data from Arduino is transferred to main clouds sensor via Wi-Fi module through Internet backbone, and when liquid comes to certain level below the overflow level, then motor is put off through wireless media. The

ultrasonic sensor helps to measure the liquid level, and when the liquid flow reaches certain level, then the liquid flow can be stopped automatically by turning the motor off or close the liquid flow in pipe by the help of Arduino. The liquid flow sensor measures the quantity of liquid flow through pipe.

For a specific amount of time in a pipe, the data will be sent to cloud for storage and analysis purposes. The other sensors like temperature, pH and turbidity sensor measure the liquid quality and help to determine whether the liquid is useful for chemical process.

48.5 Results and Discussions

48.5.1 Implementation

In this work, the liquid flow rate, current liquid flow rate and output liquid quantity are displayed in the form of graph. If the flow rate decreases, then there is a leakage in the pipe, and the liquid quantity data is also used for meter purpose. The output is shown below. Figure 48.6 illustrates sample output of liquid flow rate.

Implementation is the process of putting a decision or plan into execution and is the realization of an application, or execution of a model, design, specification, standard algorithm or policy. Such implementation is the action that must follow any preliminary thinking in order for something to actually happen. The liquid flow sensor is inserted inside the pipe, it will read the data, and the flow sensor is connected to the Arduino through wires.

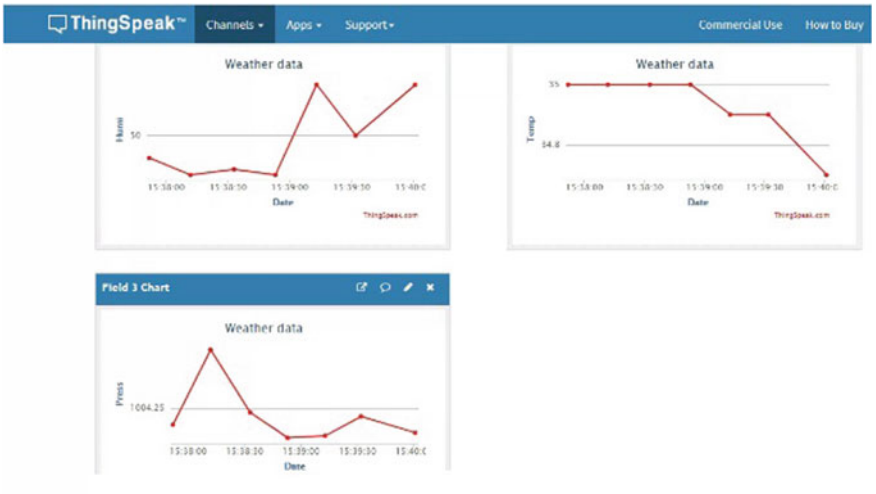


Fig. 48.6 Sample output of liquid flow rate

The Arduino is the microcontroller used in this project, and Arduino IDE is used for writing the program. It will receive the data from the sensor, and the Arduino will process the data and calculate the liquid flow rate and volume of liquid flows. The processed data is transmitted to the application through ESP8266 Wi-Fi module for displaying the result. The result is displayed in the form of graph, and it shows the flow rate and volume of liquid flows through the pipe.

48.6 Conclusion

In this paper, a design pattern of IoT-based liquid monitoring system is introduced. It is an innovative idea for automatic control of liquid flow if liquid is not in proper position in chemical plants. It is concluded that large quantity of liquid flow monitoring and its leakage is more essential in real-time work. This paper illustrates about sensor applications of IoT-based system to monitor chemical plant liquid flow attributes measurement. This can be achieved through wireless media with IoT applications. Data collection through sensors would increase the analysis of liquid flow for better solution of liquid flow problems. The data is sent to the cloud server via Wi-Fi module ESP8266. So, this application will be the best challenger in real-time monitoring and control system and is used to solve all the liquid-related problems.

References

1. Ramakala, R., Thayammal, S., Ramprakash, A., Muneeswaran, V.: Impact of ICT and IOT strategies for liquid sustainability: a case study in Rajapalayam-India. In: International. Conference Computational Intelligence and Computing Research (ICCIC 18), pp. 1–4, (2017)
2. Zidan, N., Mohammed, M., Subhi, S.: An IoT based monitoring and controlling system for liquid chlorination treatment. Proceedings of International Conference Future Networks and Distributed Systems, p. 31 (2018)
3. Ramprabu, J., Paramesh, C.: Automated sensor network for monitoring and detection of impurity in drinking liquid system. Int. J. Res. Appl. Sci. Eng. Technol. **3**, 275–280 (2015)
4. Siddula, S.S., Babu, P., Jain, P.C.: Liquid level monitoring and management of dams using IoT. In: International Conference by IEEE (2018). ISBN: 978–15090–6785–5
5. Kedia, N.: Liquid quality monitoring for rural areas-a sensor cloud based economical project. In: Published in 2015 1st International Conference on Next Generation Computing Technologies (2015)
6. Abdelhafidh, M., Fourati, M., Fourati, L.C., Abidi, A.: Remote liquid pipeline monitoring system IoT-based architecture for new industrial era 4.0. In: International Conference by IEEE (2017). ISBN: 2161–5330
7. Anjana, S., Sahana, M.N., Ankith, S., Natarajan, K., Shobha, K.R.: An IoT based 6 LoWPAN enabled experiment for liquid management. In: International Conference by IEEE (2015). ISBN: 2153–1684
8. Kafli, N., Isa, K.: Internet of Things (IoT) for measuring and monitoring sensors data of liquid surface platform. In: International Conference by IEEE (2017). ISBN: 978–1–5386–1918–6

Chapter 49
K-Hop Iterative Time Synchronization Over a Linear Connected Wireless Sensor Networks



K. L. V. Sai Prakash Sakuru and N. Bheema Rao

Abstract Clock synchronization in linearly connected wireless sensor networks over a k-hop is studied and analytical treatment given in this paper. The clock offsets for a single-hop sensor network using iterative synchronization approach over the random delay model was analyzed and extended to a two-hop sensor network. An extension of that to a k-hop sensor network shows that the error in offset estimate also increases linearly. A modified two-hop algorithm was proposed and analyzed and proved that the error in offset estimate is less than the two-hop network. The modified k-hop was analyzed based on the modified two-hop algorithm. The simulation results over uniform random delay for both two-hop and altered two-hop sensor networks presented. The proposed modified two-hop and modified k-hop achieve better precision than two-hop and k-hop algorithms.

49.1 Introduction

Time synchronization is a crucial issue in wireless sensor networks. Applications such as highway, runway, tunnel, railway line, subways, gas pipeline, water pipeline, oil pipeline, line of control, agriculture field, and cricket pitch monitoring are a few examples of linear sensor networks. Most of the applications need local clocks on the sensor nodes to be time-synchronized both locally (pairwise) and network-wide (multi-hop), with the application demand degree of precision. The locally and network-wide time synchronization protocols grouped as two parts: (i) Clocks are updated periodically and converge at a common notion of time, and (ii) clocks are updated whenever an event occurs on-demand with its neighbors.

K. L. V. Sai Prakash Sakuru (✉) · N. Bheema Rao
Department of Electronics & Communication Engineering, National Institute of Technology, Warangal, Telangana 506004, India
e-mail: sai@nitw.ac.in

N. Bheema Rao
e-mail: nbr@nitw.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_49

Most of the single-hop synchronization algorithms have analytical designs to define the accuracy aspect of the contemplated synchronization schema. In the case of multi-hop synchronization, not many analytical models are present. When a pair of nodes undergoing a synchronization process with multiple synchronization packets, the relative offset is calculated based on pairwise errors. In multi-hop synchronization, the relative offset prediction error, in general, grows linearly with the number of hops. The nodes in the multi-hop compute the synchronization time with the local values present with the intermediate nodes along the path. In the past, average-based time synchronization protocols [1–7] are proposed. The average-based protocols use pairwise packet messages exchanges and use an average common consent between them, referred to as a paired median protocol.

Time synchronization algorithms for two-hop and multi-hop proposed [8, 9] and analyzed. With a confidence level of 90% thresholds [8], the TTS precision of 0.36, 0.43, 0.52, and 0.60 s, for two, four, six, and eight hops obtained, respectively, and also NP-complete. In [1, 10–13], distributed multi-hop algorithms were proposed for network-wide synchronization. In [10] for large networks [11], Cramer-Rao lower bounds with maximum likelihood estimators [12], used energy function to minimizing it for finding clock synchronization [13] extending pairwise synchronization to cover the whole network, and [1] analyzing with Gaussian delays over the entire network. The authors in [10], use a large number of time-related constraints and use loop methods to find the lower bounds on synchronization. In [14], with reduced communication overheads and simple computation, they claim to be energy efficient. In [11], Cramer-Rao lower bounds and maximum likelihood estimators used to estimate the time synchronization error.

A spanning tree formed in [15] and a pairwise synchronization run through the branches to synchronize the complete network. It found more suitable for medium-size networks than large-size systems as it incurs a significant amount of time to sync. However, node failures and tree formation consume much energy. In [16], multi-hop time synchronization protocol over MANETs with two-phase implementation proposed beacon window and synchronization phases. In [17], the authors proposed a low energy consumption network-wide adaptive multi-hop timing synchronization algorithm.

In this paper, the authors offer a node-pair iterative time synchronization algorithm and its extension to a two-hop network. Further, the same is extended to the k-hop network and shown that the error in estimation grows linearly. A modified two-hop time synchronization algorithm proposed to improve the error performance and the same used for K-hop synchronization.

The rest of the paperwork organized as given: In Sect. 49.2, the main node-pair time synchronization algorithm explained in various steps provided. In the next, two parts 3 and 4, the two-hop and K-hop time synchronizations are analyzed. In the following Sects. 49.5 and 49.6, a modified two-hop and modified K-hop algorithms were discussed and analyzed. Finally, Sect. 49.7, the simulations results are presented and Sect 8 concludes the paper.

49.2 Iterative Clock Synchronization Algorithm

In this paper [18], we provide a node-pair iterative time synchronization algorithm. We begin by considering the case when $\alpha = 1$ the skew whence we need to determine only δ, the offset. We assume that node 0 has the correct time and is sending the timestamped packets to node 1. Let t_k be the time at which node 0 send the k-th timestamped packet and r_k be the time on the clock of node 1 when this packet is received at node 1 after random delay on the network, D_k. It is easy to see that $r_k = t_k + \delta + D_k$, i.e., $\delta = r_k - t_k - D_k$.

We assume $D_k \in [a, b]$ and $0 \leq a \leq b < \infty$, i.e., D_k are bounded random variables and that a and b are known. Hence, after receiving n timestamped packets, we see that δ should satisfy $\delta \in [r_k - t_k - b, r_k - t_k - a]$ for $1 \leq k \leq n$. This means that $\delta \in [\underline{\delta_n}, \overline{\delta_n}]$ where

$$\underline{\delta_n} = \max_{1 \leq k \leq n} \{r_k - t_k\} - b = \delta - (b - \max_{1 \leq k \leq n} \{D_k\}) \tag{49.1}$$

$$\overline{\delta_n} = \min_{1 \leq k \leq n} \{r_k - t_k\} - a = \delta * (\min_{1 \leq k \leq n} \{D_k\} - a) \tag{49.2}$$

The last equality in the above is obtained by noting that $r_k = t_k + \delta + D_k$. Let $\tilde{\delta}(n) := \overline{\delta_n} - \underline{\delta_n}$ denote the width of uncertainty in our estimate of δ after n timestamps were received. Then, we have $\tilde{\delta}(n) = (b - a) - (\max_{1 \leq k \leq n} D_k - \min_{1 \leq k \leq n} D_k)$. Note that we have not made any assumptions on the random sequence D_k of delays experienced by the timestamp packets except that they are bounded.

If the a-priori bounds are accurately selected, it is intuitively clear that max will go close to b and min will go close to a, thereby $\tilde{\delta}(n)$ converges to zero. It is clear that we have not made any assumptions on the random sequence D_n the delays experienced by the timestamped packets except that they are bounded. Thus, eventually, the offset can be estimated.

Algorithm 1 Time Synchronization Algorithm on Real Time

Require: Node $0 \geq 0$ & t =real time
Ensure: Node 1 is in Node 0 comm range
Require: bound interval $[a, b]$ corresponds to Node 0
1: **if** Node 0 **then**
2: broadcast t_n
3: **else**
4: initialize $[\underline{\delta_0}, \overline{\delta_0}] = [0, +\infty]$
5: upon receiving broadcast t_n
6: $\underline{\delta_n} = max\{t_n + \delta + a, \underline{\delta_{n-1}} + r_n - r_{n-1}\}$
7: $\overline{\delta_n} = min\{t_n + \delta + b, \overline{\delta_{n-1}} + r_n - r_{n-1}\}$
8: new uncertainty bound $[\tilde{\delta}(n) = \overline{\delta_n} - \underline{\delta_n}]$
9: **end if** $\tilde{\delta}(n) = \epsilon$

49.3 Two-Hop Synchronization

Let us consider a two-hop network as shown in Fig. 49.1, where Node 0 has the correct (true) time and is transmitting a sync packet to Node 1. Let at time t_k^0 the kth timestamped packet is transmitted from Node 0 and r_k^1 be the time on the clock of Node 1 when this packet is received at Node 1 after random delay D_k^1 on the network. The D_k^1 are i.i.d. with bounded support $[a_1, b_1]$, where the values of $a_1,\ b_1$ are assumed known at Node 1.

After receiving n timestamped packets at node 1, we have δ satisfy

$$[\underline{\delta}_n^1, \overline{\delta}_n^1] \quad for \quad 1 \leq k \leq n \tag{49.3}$$

and $\hat{\delta}_n^1 := \overline{\delta}_n^1 - \underline{\delta}_n^1$ the estimated uncertainty width of the offset δ_n^1.

Now, Node 1 sends timestamped packets to Node 2. Let t_k^1 be the time at which Node 1 sends the kth packet to Node 2, containing its then uncertainty interval $[\underline{\delta}_k^1, \overline{\delta}_k^1]$. This packet is received at Node 2 at time r_k^2, after a random delay D_k^2. The D_k^2 are assumed i.i.d. with bounded support $[a_2, b_2]$, where the values of $a_2,\ b_2$ are assumed known at Node 2.

At Node 2, we see that δ should be in the interval

$$[\underline{\delta}_k^2, \overline{\delta}_k^2] \quad for \quad 1 \leq k \leq n. \tag{49.4}$$

and $\hat{\delta}_n^2 := \overline{\delta}_n^2 - \underline{\delta}_n^2$ denotes the width of the uncertainty in our estimate of δ^2 after n timestamps were received.

Hence, after receiving n timestamped packets at node 2 from both node 0 and node 1, we have

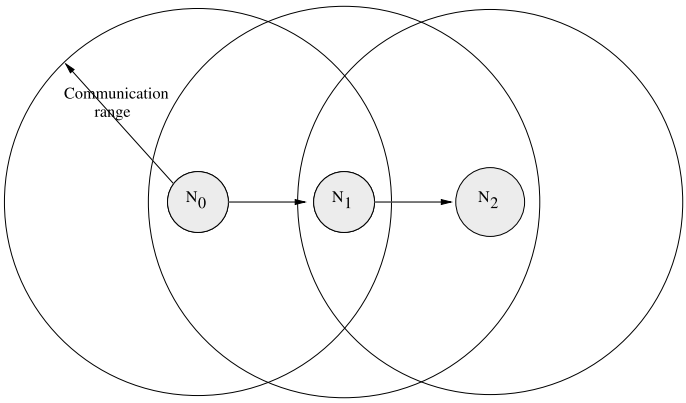


Fig. 49.1 Two-hop synchronization

$$\hat{\delta}_n^2 = \hat{\delta}_n^1 + (b_2 - a_2) - (\max_{1\le k\le n}\{D_k^2\} - \min_{1\le k\le n}\{D_k^2\}) \tag{49.5}$$

where

$$\hat{\delta}_n^1 = (b_1 - a_1) - (\max_{1\le k\le n}\{D_k^1\} - \min_{1\le k\le n}\{D_k^1\}) \tag{49.6}$$

If we look at Eq. 49.5, the uncertainty interval at the nodes 2 can be rewritten as, under the assumption that the delay bounds between all hop are same and that $a,\ b$ are known at nodes 1 and 2, then we can rewrite it as

$$\hat{\delta}_n^2 = 2(b-a) - \sum_{i=1}^{2}(\max_{1\le k\le n}\{D_k^i\} - \min_{1\le k\le n}\{D_k^i\}) \tag{49.7}$$

Hence, we can see the uncertainty interval propagates linearly, and it is almost 2 em times at the end of two-hop network.

49.4 K-Hop Synchronization

Let us consider a K-hop network where node $K-1$ will transmit a sync packet at time t_k^{K-1} the kth timestamped packet, and r_k^K be the time on the clock of node K when this packet is received at node K after random delay D_k^K on the network along with the then uncertainty interval $[\underline{\delta}_k^{K-1}, \overline{\delta}_k^{K-1}]$ at node $K-1$. The $D_{(k)}^K$ are i.i.d. with bounded support $[a_K, b_K]$, where the values of $a_K,\ b_K$ are assumed known at node K.

Proceeding on the same lines of two-hop synchronization as discussed above, the uncertainty width of clock at node K is given by

$$\hat{\delta}_n^K = \hat{\delta}_n^{K-1} + (b_K - a_K) - (\max_{1\le k\le n}\{D_k^K\} - \min_{1\le k\le n}\{D_k^K\}) \tag{49.8}$$

where

$$\begin{aligned}\hat{\delta}_n^{K-1} &= \hat{\delta}_n^{K-2} + (b_{K-1} - a_{K-1}) \\ &-(\max_{1\le k\le n}\{D_k^{K-1}\} - \min_{1\le k\le n}\{D_k^{K-1}\})\end{aligned} \tag{49.9}$$

$$\vdots$$

$$\hat{\delta}_n^1 = (b_1 - a_1) - (\max_{1\le k\le n}\{D_k^1\} - \min_{1\le k\le n}\{D_k^1\}) \tag{49.10}$$

If we look at Eq. 49.8, the uncertainty interval at the Kth hop can be rewritten as, under the assumption that the delay bounds between all hop are same and that a, b are known at all the $K-1$ nodes, then we can rewrite it as

$$\hat{\delta}_n^K = K(b-a) - \sum_{i=1}^{K}(\max_{1\le k\le n}\{D_k^i\} - \min_{1\le k\le n}\{D_k^i\}) \tag{49.11}$$

Hence, we can see the uncertainty interval propagates linearly, and it is almost K times at the end of $K-$ hop network.

49.5 Proposed Algorithm—Modified Two-Hop Synchronization

The main idea is to capture the convergence at Node 2 when it gets alternative packets from the Node 1 which is in the processes of convergence and also from Node 0 which is transmitting the synchronization packets with Node 1 as simply packet forwarder under the assumption it takes zero delay in forwarding. Let Node 1 has a support $[a_1, b_1]$ and has an uncertainty width $\hat{\delta}_i^1$ which is represented by $[\underline{\delta}_i^1, \overline{\delta}_i^1]$ to which it has converge after i^{th} iteration is transmitted to Node 2 as shown in Fig. 49.2. Now, let Node 2 has a support $[a_2, b_2]$ uses the information transmitted by Node 1 and converges to an uncertainty width $\hat{\delta}_i^2$ which is represented by $[\underline{\delta}_i^2, \overline{\delta}_i^2]$.
The analysis is divided into four parts:

- Node 0 and Node 1 synchronization.
- Node 0 and Node 2 synchronization (with Node 1 as simple relay)
- Node 1 and Node 2 synchronization
- Node 2 synchronization with help of above two packets

Node 0 and Node 1 synchronization The uncertainty width is given by:

$$\hat{\delta}_n^1 = \overline{\delta}_n^1 - \underline{\delta}_n^1 = (b_1 - a_1) - (\max_{1\le i\le n}\{D_i^1\} - \min_{1\le i\le n}\{D_i^1\}) \tag{49.12}$$

Node 0 and Node 2 synchronization (with Node 1 as simple relay) The uncertainty width is given by:

$$\begin{aligned}\hat{\delta}_n^2 = \overline{\delta}_n^2 - \underline{\delta}_n^2 = (b_1 - a_1) + (b_2 - a_2)\\ -(\max_{1\le i\le n}\{D_i^2 + D_i^1\} - \min_{1\le i\le n}\{D_i^2 + D_i^1\})\end{aligned} \tag{49.13}$$

Node 1 and Node 2 synchronization The uncertainty width is given by:

$$\hat{\delta}_n^{\overline{2}} = \overline{\delta}_n^{\overline{2}} - \underline{\delta}_n^{\overline{2}} = (b_2 - a_2) - (\max_{1\le i\le n}\{D_i^{\overline{2}}\} - \min_{1\le i\le n}\{D_i^{\overline{2}}\}) \tag{49.14}$$

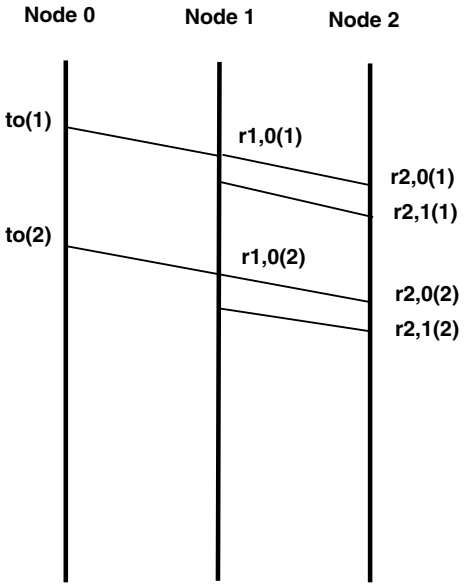


Fig. 49.2 Modified two-hop synchronization

Node 2 synchronization with help of above two packets (one coming directly via Node 1 and the other from Node 1 while synchronizing) The uncertainty width of clock at Node 2 is given by

$$\hat{\delta}_n^{\hat{2}} = (b_1 - a_1) + (b_2 - a_2) - \max\{\max_{1\le i\le n}\{D_i^1 + D_i^2\}, \max_{1\le i\le n}\{D_i^1\} + \max_{1\le i\le n}\{D_i^{\hat{2}}\}\} + \min\{\min_{1\le i\le n}\{D_i^1 + D_i^2\}, \min_{1\le i\le n}\{D_i^1\} + \min_{1\le i\le n}\{D_i^{\hat{2}}\}\} \quad (49.15)$$

49.6 Proposed Algorithm—Modified K-Hop Synchronization

Modified k-hop synchronization The uncertainty width of clock at Node K is given by

$$\hat{\delta}_n^{\hat{k}} = \sum_{i=1}^{K}(b_K - a_K) - \max\{\max_{1\le i\le n}\{D_i^1 + D_i^2 + \cdots + D_i^K\}, \ldots,$$
$$\max_{1\le i\le n}\{D_i^1\} + \max_{1\le i\le n}\{D_i^{\hat{2}}\} + \cdots + \max_{1\le i\le n}\{D_i^{\hat{K}}\}\}$$
$$+ \min\{\min_{1\le i\le n}\{D_i^1 + D_i^2 + \cdots + D_i^K\}, \ldots,$$
$$\min_{1\le i\le n}\{D_i^1\} + \min_{1\le i\le n}\{D_i^{\hat{2}}\} + \cdots + \min_{1\le i\le n}\{D_i^{\hat{K}}\}\} \quad (49.16)$$

49.7 Simulations

The two-hop and modified two-hop iterative time synchronization algorithm is analytically derived and simulated in MATLAB. The a-priori uncertainty bounds are considered a 1-time unit as the lower bound and 5-time units as the upper bound. Node 0 transmits timestamped message packets periodically with a regular interval of unit time. On the reception of the message packet at node 1, it is forwarded to node 2. The process is monitored for the behavior of the uncertainty width, with incremental message transmission packets. The independent and identically distributed random delays assumed to take uniform density function with the above uncertainty support. The simulation studies carried out considering the distributions of $D^1 \& D^2$ to be uniform and bounded with a_1 =1 unit time, b_1 = 5 unit time, a_2 = 1 unit time and b_2 = 5 unit time and measurements made over 25 packets, as shown in Fig. 49.3.

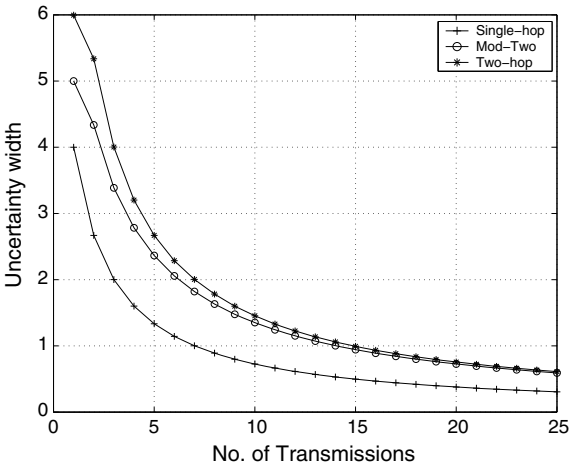


Fig. 49.3 Expectation of $\tilde{\delta}_n^2$ for uniform distributions

49.8 Conclusions and Future Work

An iterative time synchronization one-hop algorithm proposed and analytically derived. A two-hop synchronization algorithm was calculated based on the said one-hop model and obtained the analytical expression. In similar lines, for linear wireless sensor networks, a k-hop synchronization was proposed, and an analytical expression was derived, showing that the uncertainty error propagates linearly. We proposed a modified two-hop synchronization algorithm to combat the error propagation and given a systematic treatment. The same procedure extended for the k-hop network. From Fig. 49.3, we can conclude that if the data of the uncertainty width of the interval is carefully selected, it will impact the convergence curve. The simulation results show that the uncertainty error is almost double in the case of a two-hop network compared to a hop network. The proposed modified two-hop system has better uncertainty width compared to a two-hop network. Hence, as the number of hops increases, the error will not grow linearly. The results are encouraging for large systems, and looking further for mid-course correction strategies to cover the entire network.

References

1. Gang, X., Shalinee, K.: Analysis of distributed consensus time synchronization with gaussian delay over wireless sensor networks. EURASIP J. Wireless Commun. (2009)
2. Sommer, P., Wattenhofer, R.: Gradient clock synchronization in wire- less sensor networks. In: IPSN'09, April 13–16 2009, San Francisco, CA, USA (2009)
3. Freris, N., Graham, S., Kumar, P.: Fundamental limits on synchronizing clocks over networks. IEEE Trans. Automatic Control **56**(6), 1352–1364 (2011)
4. Leng, M., Wu, Y.C.: Distributed clock synchronization for wireless sensor networks using belief propagation. IEEE Trans. Signal Process. **59**(11), 5404–5414 (2011)
5. Djenouri, D., Merabtine, N., Mekahlia, F.Z., Doudou, M.: Fast distributed multi-hop relative time synchronization protocol and estimators for wireless sensor networks. Ad Hoc Netw. **11**(8), 2329–2344 (2013)
6. Li, S.Z.: Markov Random Field Modeling in Image Analysis, 3rd edn. Springer, Heidelberg (2009)
7. Sundararaman, B., Buy, U., Kshemkalyani, A.D.: Clock synchronization for wireless sensor networks: a survey. AdHoc Networks **3**(3), 281–323 (2005)
8. Wang, J., Zhang, S., Gao, D., Wang, Y.: Two-hop time synchronization protocol for sensor networks. EURASIP J. Wireless Commun. Network. 2014:39 (2014). http://jwcn.eurasipjournals.com/content/2014/1/39
9. Cui, L.Z., Yuan, M.M., Li, Z., et al.: Research and implementation of low power two-way WSN time synchronization algorithm for coal mine. Chin. J. Sens. Actuat. **27**(9), 1253–1259 (2014)
10. Solis, R., Borkar, V.S., Kumar, P.R.: A new distributed time synchronization protocol for multihop wireless networks. In: Proceedings of the 45th IEEE Conference on Decision and Control. https://doi.org/10.1109/CDC.2006.377675
11. Djenouria, D., Merabtineb, N., Mekahliab, F.Z., Doudoua, M.: Fast distributed multi-hop relative time synchronization protocol and estimators for wireless sensor networks. Elsevier Ad Hoc Networks (2013). https://doi.org/10.1016/j.adhoc.2013.06.001

12. Chen, W., Jin, X., Li, Z., Zhang, X., Hong, L.: Clock synchronization for distributed multi-hop wireless networks using Markov Random field. J. Phys.: Conf. Ser. **1087**, 052008 (2018). https://doi.org/10.1088/1742-6596/1087/5/052008
13. Cheng, K.Y., Lui, K.S., Wu, Y.C., et al.: A distributed multihop time synchronization protocol for wireless sensor networks using pairwise broadcast synchronization. IEEE T Wireless Commun. **8**(4), 1764–1772 (2009)
14. Kaseva, V., Hämäläinen, T.D., Hännikäinen, M.: Time synchronization for resource-constrained multi-hop wireless sensor networks based on hop delay estimation. In: SENSORCOMM 2011, ISBN: 978-1-61208-144-1
15. Maroti, M., Kusy, B., Simon, G., Ledeczi, A.: Robust multi-hop time synchronization in sensor networks. In: ICWN '04, pp. 454–460 (2004)
16. Chen, G.-N., Wang, C.-Y.,Hwang, R.-H.: MTSP: Multi-hop time synchronization protocol for IEEE 802.11 wireless Ad Hoc network. In: WASA 2006: Wireless Algorithms, Systems, and Applications. pp. 664–675
17. Noh, K.-L., Serpedin, E.: Adaptive multi-hop timing synchronization for wireless sensor networks. In: ISSPA (2007). https://doi.org/10.1109/ISSPA.2007.4555613
18. Sai Prakash Sakuru, K.L.V., Bheema Rao, N.: An iterative node-pair time synchronization (INTS) for wireless sensor networks. In: ESIC 2020, March 02–04, 2020. Yupia, Arunachal Pradesh, India (2020)

Chapter 50
A Convolutional Neural Network Architecture for Tomato Leaf Disease Detection Using Data Augmentation



Matta Bharati Devi and K. Amarendra

Abstract Convolutional neural networks have set state-of-the-art results for many challenging problems in the field of computer vision. In this article, we design and implement a six-layered convolutional neural network for tomato leaf disease detection. Our model is trained on tomato plant leaf images of ten different classes representing various diseases of tomato plant. To increase the model generalizability, we employed various data augmentation techniques and increased the size of training data. Batch normalization layers after every convolution and dropout induced after fully connected layers made the model immune from overfitting. Our proposed model is able to outperform various existing works on tomato leaf disease detection and achieved an accuracy of 92.3% on test data.

50.1 Introduction

Plant leaf disease identification is one of the major challenges faced by farmers in agriculture. It is very important to correctly identify the type of leaf disease for appropriate use of pesticides. Any mistakes in accurate identification of plant leaf diseases leads to reduced yield. According to statistics, 52% of the Indian population depends on agriculture. Most of the local farmers and gardeners grow tomatoes in their farms and gardens as it is one of the most required vegetables in our day-to-day life. But they are unable to find appropriate pesticides to use when they are infected by different types of pests and diseases. The reason behind this is tomato plants get affected by various diseases in its life span of being productive, and identifying each disease accurately is a challenge. Hence, there is a requirement to automate the process of tomato leaf disease detection. In recent years, tomato leaf disease detection has been an active area of research, and several methods were proposed

M. Bharati Devi · K. Amarendra (✉)
Department of CSE, Koneru Lakshmaiah Education Foundation, Vijayawada, Andhra Pradesh, India
e-mail: amarendra@kluniversity.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_50

to automate this process using several conventional machine learning approaches [1, 2].

Feature extraction techniques like histogram of oriented gradients (HOG), local binary patterns (LBP) and grey-level co-occurrence matrix (GLCM) are used in literature to extract texture-based features from plant leaf images [3]. These features were fed to popular classifiers like support vector machine (SVM) to categorize different types of diseased leaves [4]. However, this type of feature engineering requires sound domain expertise to extract discriminative information from leaf images. With the use of deep learning models like convolutional neural networks (CNNs), there is a great process in this area of research. Most of them designed their own CNN architectures to tackle this problem of tomato leaf disease detection [5]. But, large amount of training data is required to train these models [6, 7]. To accomplish this problem, various pre-trained models trained on large amount of labelled data were introduced. Xception pre-trained network was used to detect and classify tomato leaf diseased images [8, 9]. VGG16 and ResNet models were fused, and resultant model was trained on leaf images of various plants to detect diseased leaves [10]. All the models have large number of layers and need huge computing resources for training.

In this article, we use several data augmentation techniques to increase the size of training data and design a simple convolutional neural network architecture with batch normalization layers. We train our proposed architecture using tomato leaves of ten different classes taken from plant village dataset. Our proposed model achieves competitive results with existing pre-trained models for tomato leaf disease prediction. Our model has a smaller number of trainable parameters when compared to existing networks.

50.2 Related Work

This section provides discussion of various approaches used for plant leaf disease detection. Accurate diagnosis of plant leaf disease has been an important area of research in the field of machine learning. Most of the previous works concentrated on using image processing and machine learning-based techniques to detect the type of plant disease from leaf images. A K-nearest neighbour (K-NN) classifier with grey-level co-occurrence matrix (GLCM) texture features of plant leaf images was used to identify plant leaf diseases [11, 12]. Another machine learning-based system was proposed for grapes plant leaf disease detection by Harshal Waghmare et al. First, background of all images was removed, and segmentation is performed as a preprocessing step. A high-pass filter is applied on segmented images to analyse disease part of the leaf. Local binary patterns (LBP)-based texture features are extracted from preprocessed images, and these features were used to identify different types of grape plant diseases using support vector machine (SVM) classifier [13]. A cotton leaf disease detection and classification technique based on machine learning and image processing tools is proposed by V Pooja et al. Initially, region of interest (ROI) is segmented from plant leaf images using image processing tools, and features are

extracted. These features were passed to SVM classifier to identify the type of disease [14, 15].

In recent years, convolutional neural networks (CNN) marked breakthrough results for various tasks like object recognition, image classification and medical image diagnosis [16, 17] tasks. Different CNN architectures are built for plant leaf disease detection in literature. A deep CNN architecture is designed and trained on plant village dataset to identify 38 different types of diseases in 14 types of plant species [18, 19]. Another deep CNN architecture is proposed for detecting apple plant leaf diseases in real time [20]. Gandhi et al. designed and trained a deep CNN model on augmented data generated using generative adversarial networks (GANs) [21, 22]. A global pooling dilated convolutional neural network (GPDCNN) was employed to detect cucumber leaf diseases by blending global pooling with dilated convolution [23, 24]. Most of these models proved the efficiency of CNNs for extracting discriminant features from plant leaf images for disease identification over conventional machine learning approaches. But, these deep CNNs require large amount of labelled data for training. To overcome this issue, transfer learning was introduced.

Transfer learning is a deep learning-based paradigm, where the knowledge acquired by a model trained on large labelled data is transferred to another model to perform similar task. This is accomplished using pre-trained deep CNNs [25]. Pre-trained models are those, which are trained on ImageNet dataset containing millions of labelled images belonging to 1000 different classes [26]. AlexNet, VGG Net, GoogLeNet, ResNet and NASNet are some deep CNN architectures trained on ImageNet dataset. Even though they are trained on ImageNet dataset, these pre-trained models worked well for medical image classification tasks [27]. LeNet deep CNN architecture is trained on 18,160 tomato plant leaf images taken from plant village dataset and resulted in 94.5% test accuracy [28]. Nithish Kannan et al. used ResNet pre-trained model with 50 layers for tomato leaf disease prediction [29]. When two different deep CNN models, AlexNet and Squeeze Net, are fine-tuned on same tomato leaf dataset taken from plant village, it is observed that AlexNet achieved the highest accuracy of 95.65% and SqueezeNet obtained accuracy 94.3% [30].

Most of these models obtained better performance. In this work, we propose a deep convolutional neural network architecture whose performance is competitive with the performance of these pre-trained models.

50.3 Methodology

Our proposed method for tomato plant leaf disease prediction consists of two different phases. They are preprocessing, model training and evaluation.



Fig. 50.1 Background removal from tomato leaf

50.3.1 Data Preprocessing

Labelled tomato plant leaf images were taken from plant village dataset. These images have little amount of background variations. So, leaf regions were segmented from the images by applying appropriate masks on images. This helps the model to learn necessary features without any variations. Figure 50.1 represents background subtracted tomato plant leaf image and appropriate mask used for it.

After performing segmentation operation, all the resultant images are reshaped to a fixed dimension of $128 \times 128 \times 3$ to increase the speed of training process. After reshaping, all images are normalized to bring all pixel intensity values into common range by computing mean and standard deviation of pixels. This type of normalization is regarded as Z-score normalization.

50.3.2 Data Augmentation

To increase the size of training data, we augment data using several data augmentation operations like rotation, shearing, flipping and zooming.

- Rotation: Original leaf image is rotated by 'x' degrees.
- Shearing: Different orientations of tomato leaf images are created by applying shear transformation.
- Flipping: Leaf images are flipped either horizontally or vertically.
- Zooming: Zoom in or zoom out operations create additional labelled images.
- Brightness Adjustment: Original image replications are created by different variations in brightness levels.

Data augmentation techniques make the model more robust by creating different variations in training data. Figure 50.2 represents invariant images created by these augmentation techniques.

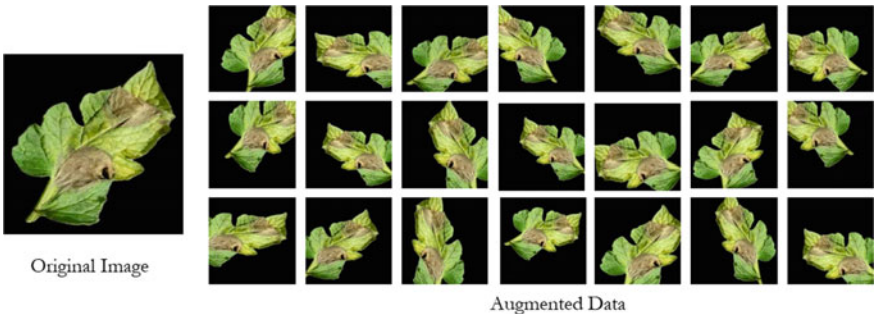


Fig. 50.2 Sample images of augmented data

50.3.3 Convolutional Neural Network for Tomato Leaf Disease Identification

We design a deep convolutional neural network architecture and train it using preprocessed tomato leaf images. Proposed CNN architecture consists of four convolution layers, each one followed by a batch normalization and max pooling layers. We used (3×3) kernels for all convolution and a (2×2) window for max pooling operations. By rigorous experiments, we decided the number of kernels used for each convolution operation. First convolution layers use 32 kernels, and next three convolution layers use 64 kernels. Rectified linear unit (ReLU) activation function is applied on each convoluted feature map. We use 'same' padding to retain the information from convolutional feature maps. All the convolution feature maps are after fourth convolutional layers which are passed to flatten layer to vectorize them.

A dense layer with 64 units and ReLU activation function is followed after flatten layer which learns discriminative information obtained from feature maps generated by convolutional layers. A dropout layer with 0.5 rate is added after dense layer to provide regularization and prevent model from overfitting. Finally, another dense layer with 10 units and softmax activation function is added at the end of network for classification. Figure 50.3 represents architecture of proposed deep CNN architecture. Our CNN architecture is simple in terms of number of parameters when

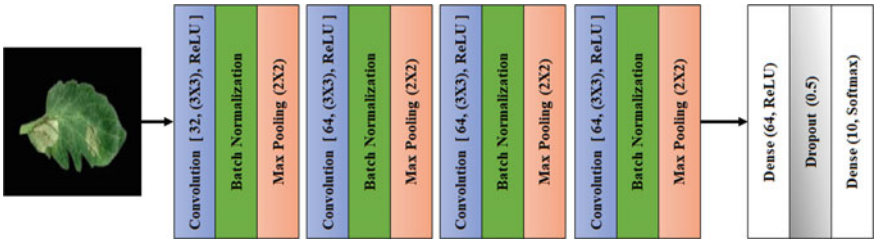


Fig. 50.3 Proposed deep CNN architecture

compared with other deep CNN architectures designed for tomato leaf disease identification. Since model is trained on augmented data, model learns invariant features and results in improved performance.

50.3.4 Training Details

Proposed CNN architecture is trained on tomato leaf disease dataset. Optimal values of kernels learned using categorical cross-entropy loss function and error are optimized by ADAM optimizer. Data batch of size 64 is loaded every time during training to reduce memory burden on system. Model is executed for 50 epochs, during training best weights with minimal validation loss are saved using model checkpointing mechanism.

50.4 Experimental Results

This section provides clear picture of experiments conducted and results obtained using proposed method. First overview of dataset used for experiments is described followed by evaluation metrics used to measure the performance of proposed deep CNN architecture. Finally, a summary of experiments and their results are provided.

50.4.1 Tomato Leaf Disease Dataset

We used tomato leaf disease dataset, part of plant village dataset which contains 54,306 samples of 26 type of diseased leaf images belonging to 14 types of plant species. This dataset contains 18,160 samples of tomato plant representing both healthy and diseased leaves. Diseased leaves are broadly labelled to nine different types of diseases. Totally, 14,530 samples are considered for training, and 3630 samples are used for testing model performance. Figure 50.4 represents sample images of diseased tomato leaves taken from plant village dataset (Table 50.1).

50.4.2 Performance Evaluation Measures

Different classification model performance evaluation measures like accuracy, precision, recall and F1-score are calculated to prove the efficiency of proposed deep CNN network on test data. These measures can be computed using confusion matrix.

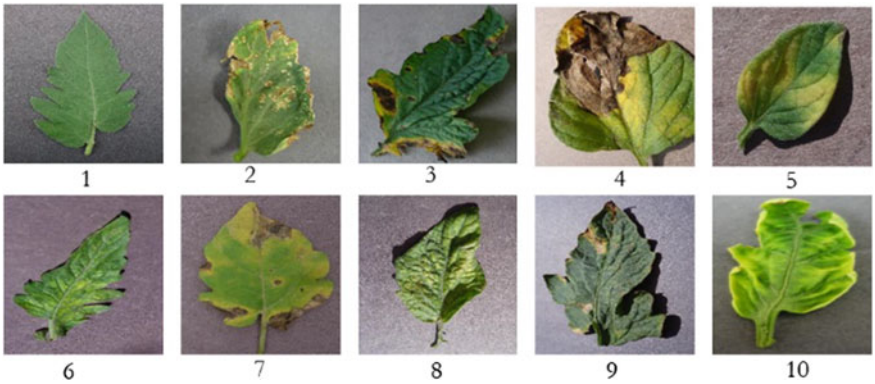


Fig. 50.4 Samples of diseased tomato leaves from tomato leaf disease dataset

Table 50.1 Dataset summary for tomato leaf disease dataset

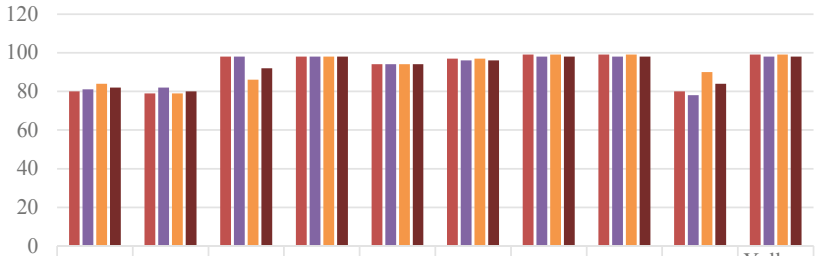
Category	# Train	# Test
Healthy	1273	318
Bacterial spot	1702	425
Early blight	800	200
Late blight	1527	382
Leaf mould	762	190
Mosaic virus	298	75
Septoria leaf spot	1417	354
Spider mites	1342	334
Target spot	1123	281
Yellow leaf curl virus	4286	1071
Total	14,530	3630

50.4.3 Results Analysis

We trained our proposed CNN with preprocessed tomato leaf images belonging to ten different classes. Performance of our CNN is picturized in Fig. 50.5, in terms of accuracy precision, recall and F1-score.

From Fig. 50.5, it is clear that proposed model is able to correctly classify tomato leaf images belonging to early blight, late blight, mosaic virus, septorial leaf spot, spider mites and yellow leaf curl virus classes. Due to a smaller number of training samples for healthy, bacterial spot and target spot classes, there is a bit reduced accuracy for these classes.

Our proposed CNN architecture is able to outperform a ToLeD CNN architecture proposed in [5]. From Table 50.2, it is clear that the number of parameters is less



	Healthy	Bacterial Spot	Early Blight	Late Blight	Leaf Mold	Mosaic Virus	Septorial Leaf Spot	Spider Mites	Target Spot	Yellow Leaf Curl Virus
Accuracy	80	79	98	98	94	97	99	99	80	99
Precision	81	82	98	98	94	96	98	98	78	98
Recall	84	79	86	98	94	97	99	99	90	99
F1 Score	82	80	92	98	94	96	98	98	84	98

Fig. 50.5 Performance of proposed CNN in terms of accuracy, precision, recall and F1-score

Table 50.2 Performance comparison of proposed CNN with existing models

Model	Accuracy (%)	# Parameters
VGG16 [4]	77.2	9,449,482
MobileNet [4]	63.75	20,507,146
Inception [4]	63.4	41,247,146
ToLeD—CNN [4]	91.2	208,802
Proposed CNN	92.3	241,866

when compared with pre-trained networks like VGG16, MobileNet and Inception. Our model is superior in terms of accuracy and also number of parameters.

50.5 Conclusion

Plant leaf disease detection is one of the major challenges faced by farmers in agriculture. It is very important to correctly identify the type of leaf disease for appropriate use of pesticides. Any mistakes in accurate identification of leaf diseases lead to reduced yield. With the help of proposed deep CNN architecture, tomato leaf diseases can be correctly identified. When compared to existing methods for tomato leaf disease identification, our approach is simple and robust in terms of parameters. This study can be extended to segment diseased portions of leaf for further improvement of accuracy.

References

1. Patil, R., Kumar, S.: A Bibliometric survey on the diagnosis of plant leaf diseases using artificial intelligence (2020)
2. Dudi, B., Rajesh, V.: Medicinal plant recognition based on CNN and machine learning. Int. J. Adv. Trends Computer Sci. Eng. **8**(4), 628–631 (2019)
3. Mahapatra, S., Kannoth, S., Chiliveri, R., Dhannawat, R.: plant leaf classification and disease recognition using SVM, a machine learning approach. Sustain. Humanosphere **16**(1), 1817–1825 (2020)
4. Bhagat, M., Kumar, D., Haque, I., Munda, H.S., Bhagat, R.: Plant leaf disease classification using grid search based SVM. In: 2nd International Conference on Data, Engineering and Applications (IDEA), pp. 1–6. IEEE (2020)
5. Agarwal, M., Singh, A., Arjaria, S., Sinha, A., Gupta, S.: ToLeD: Tomato leaf disease detection using convolution neural network. Procedia Computer Sci. **167**, 293–301 (2020)
6. Xie, X., Ma, Y., Liu, B., He, J., Li, S., Wang, H.: A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks. Front. Plant Sci. **11** (2020)
7. Balakrishna, G., Moparthi, N.R.: Study report on Indian agriculture with IoT. Int. J. Electr. Computer Eng. **10**(3), 2322 (2020)
8. Pattnaik, G., Shrivastava, V.K., Parvathi, K.: Transfer learning-based framework for classification of pest in tomato plants. Appl. Artif. Intell. 1–13 (2020)
9. NagaGeetha, M., Ramesh, N.V.K.: An efficient IoT based smart irrigation system and plant diseases detection: a review. Int. J. Eng. Technol **7**(2.7), 661–664 (2018)
10. Kumar, A., Razi, R., Singh, A., Das, H.: Res-VGG: A novel model for plant disease detection by fusing VGG16 and ResNet models. In: International Conference on Machine Learning, Image Processing, Network Security and Data Sciences, pp. 383–400. Springer, Singapore (2020)
11. Trivedi, J., Shamnani, Y., Gajjar, R.: Plant leaf disease detection using machine learning. In: International Conference on emerging technology trends in electronics communication and networking, pp. 267–276. Springer, Singapore (2020)
12. Kalavala, S.S., Sakhamuri, S., Prasad, B.B.V.S.V.: An efficient classification model for plant disease detection. Int. J. Innov. Technol. Exploring Eng. **8**(7), 126–129 (2019)
13. Waghmare, H., Kokare, R., Dandawate, Y.: Detection and classification of diseases of grape plant using opposite colour local binary pattern feature and machine learning for automated decision support system. In: 2016 3rd International Conference on Signal Processing and Integrated Networks (SPIN), pp. 513–518. IEEE (2016)
14. Pooja, V., Das, R., Kanchana, V.: Identification of plant leaf diseases using image processing techniques. In: 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR), pp. 130–133. IEEE (2017)
15. Balram, G., Kumar, K.K.: Smart farming: disease detection in crops. Int. J. Eng. Technol. **7**(2.7), 33–36 (2018)
16. Bodapati, J.D., Veeranjaneyulu, N., Shareef, S.N., Hakak, S., Bilal, M., Maddikunta, P.K.R., Jo, O.: Blended Multi-modal deep ConvNet features for diabetic retinopathy severity prediction. Electronics **9**(6), 914 (2020)
17. Balaji Bhanu, B., Hussain, M.A., Prasad, A., Mirza, M.A.: Exploration of crop production improvement through various agriculture monitoring systems. Int. J. Innov. Technol. Exploring Eng. **8**(11), 3747–3750 (2019)
18. Trivedi, J., Shamnani, Y., Gajjar, R.: Plant leaf disease detection using machine learning. In: Gupta S., Sarvaiya J. (eds.) Emerging Technology Trends in Electronics, Communication and Networking (ET2ECN 2020). Communications in Computer and Information Science, vol. 1214. Springer, Singapore (2020)
19. Anusha, A., Guptha, A., Rao, G.S., Tenali, R.K.: A model for smart agriculture using IOT. Int. J. Innov. Technol. Exploring Eng. **8**(6), 1656–1659 (2019)
20. Jiang, P., Chen, Y., Liu, B., He, D., Liang, C.: Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks. IEEE Access **7**, 59069–59080 (2019)

21. Gandhi, R., Nimbalkar, S., Yelamanchili, N., Ponkshe, S.: Plant disease detection using CNNs and GANs as an augmentative approach. In: 2018 IEEE International Conference on Innovative Research and Development (ICIRD), pp. 1–5. IEEE (2018)
22. Inthiyaz, S., Prasad, M.V.D., Lakshmi, R.U.S., Sai, N.S., Kumar, P.P., Ahammad, S.H.: Agriculture based plant leaf health assessment tool: a deep learning perspective. Int. J. Emerg. Trends Eng. Res. **7**(11), 690–694 (2019)
23. Zhang, S., Zhang, S., Zhang, C., Wang, X., Shi, Y.: Cucumber leaf disease identification with global pooling dilated convolutional neural network. Comput. Electron. Agric. **162**, 422–430 (2019)
24. Sai Prasanth,P.V.V., Veera Prasad, G., Krian Babu, M.: Automatic irrigation system using Arduino uno. Int. J. Eng. Adv. Technol. **8**(5), 457–459 (2019)
25. Bodapati, J.D., Veeranjaneyulu, N.: Abnormal network traffic detection using support vector data description. In: Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications, pp. 497–506. Springer, Singapore (2017)
26. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009).
27. Dondeti, V., Bodapati, J.D., Shareef, S.N., Naralasetti, V.: Deep convolution features in non-linear embedding space for fundus image classification deep convolution features in non-linear embedding space for fundus image classification
28. Tm, P., Pranathi, A., SaiAshritha, K., Chittaragi, N.B., Koolagudi, S.G.: Tomato leaf disease detection using convolutional neural networks. In: 2018 Eleventh International Conference on Contemporary Computing (IC3), pp. 1–5. IEEE (2018)
29. Kaushik, M., Prakash, P., Ajay, R., Veni, S.: Tomato leaf disease detection using convolutional neural network with data augmentation. In: 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 1125–1132. IEEE (2020)
30. Durmuş, H., Güneş, E.O., Kırcı, M.: Disease detection on the leaves of the tomato plants by using deep learning. In: 2017 6th International Conference on Agro-Geoinformatics, pp. 1–5. IEEE (2017)

Chapter 51
Development of Robust Framework for Automatic Segmentation of Brain MRI Images



K. Bhima and A. Jagan

Abstract The robust framework for automatic segmentation of brain MRI images is developed in this work. The brain MRI image has been typically used to analyzing the brain anatomical structures and disorder in the brain. The identification of viable segmentation framework for brain MRI image is notoriously difficult problem. The major difficulty in brain MRI image analysis is mostly due to the unattainable of efficient segmentation framework. The proposed work develops a robust framework for classification and segmentation of brain MRI images. The developed robust framework consists of two phases, i.e., classification of input MRI images and detection of tumor. The automatic SVM classification and improved watershed methods are used in the robust framework to improve the segmentation process and improve the segmentation accuracy. In the first phase, MRI images classification is obtained with training and testing process with SVM classifier. In the second phase, segmentation of tumor region in MRI image is presented with improved watershed method. The developed robust framework is quantitatively evaluated on benchmark brain MRI image dataset.

51.1 Introduction

The magnetic resonance imaging (MRI) [1–3] is primary brain imaging modalities to capture and produce a computerized image for diagnosis and future treatment planning. MRI image does not emit ionic radiation like other modalities X-rays and computer tomography (CT) scan. The manual identification of tumor in MRI image is the time-consuming and tedious process. Since the numbers of patients are increasing, the acquired number of images also increases. Brain MRI image is characteristically used to capture images in diverse modalities [4–6]. Hence, there is a requirement of

K. Bhima (✉) · A. Jagan
B.V. Raju Institute of Technology, Narsapur, Telangana, India

A. Jagan
e-mail: jagan.amgoth@bvrit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_51

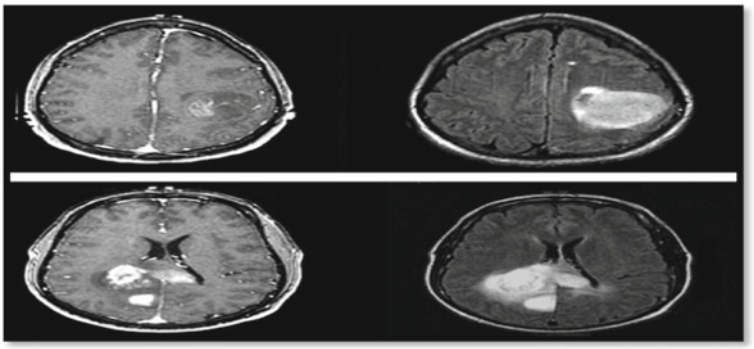


Fig. 51.1 Demonstration of four patients brain MRI Images

robust automatic segmentation framework that is competent to process and generate accurate segmentation results. The MRI or the magnetic resonance imaging images are a powerful brain imaging technique which is used to capture the brain cranial structure and produce a computerized image for further analysis and diagnosis of the brain tumor.

The output format of MRI technique as shown in Fig. 51.1 depends on the signal of the magnetic field and how it deteriorates commonly known as relaxation. The relaxation is the pure reflection of the magnetic resonance. Bilateral filter [4, 7, 8] is used for preprocessing of input MRI images. In this research work, the most popular SVM method [6, 9, 10] is used for classification of input brain MRI images into tumor and non-tumor MRI images, and the most accepted Watershed method [4, 5, 11, 12] is used for segmentation and detection of brain tumor in MRI images.

51.2 Development of Robust Framework

The proposed work develops a fully automatic framework for brain tumor detection and segmentation in brain MRI images. The proposed method consists of two phases, i.e., classification of input MRI images and brain tumor region extraction. The proposed framework has an automatic MRI image classification followed by tumor segmentation. In the first phase, MRI image classification is done through training and testing process using SVM classifiers. The infinite feature selection (IFS) method is used to detect optimal features in input MRI images. The SVM classification is used to automatically classify the input MRI image into a normal or tumor MRI image and identify the tumor. The improved watershed method is used in phase-II to extract the tumor in abnormal tumor MRI images that is identified in phase-I.

Phase-I: The proposed method has an automatic classification of brain MRI images into tumor and non-tumor MRI images.

(a) The MRI classification is done through training and testing process using SVM classifier with 8 × 8 patches.
(b) Three optimal features are chosen using infinite feature selection (IFS) method.
(c) The purpose of the first phase is to automatically cluster the input MRI image into a normal or tumor MRI image and identify the tumor.

Figure 51.2 represents the robust framework for automatic classification of brain MRI images and segmentation of tumor in MRI Images.

Phase-I: The Classification of Tumors and Non-tumors MRI images with SVM classifier.

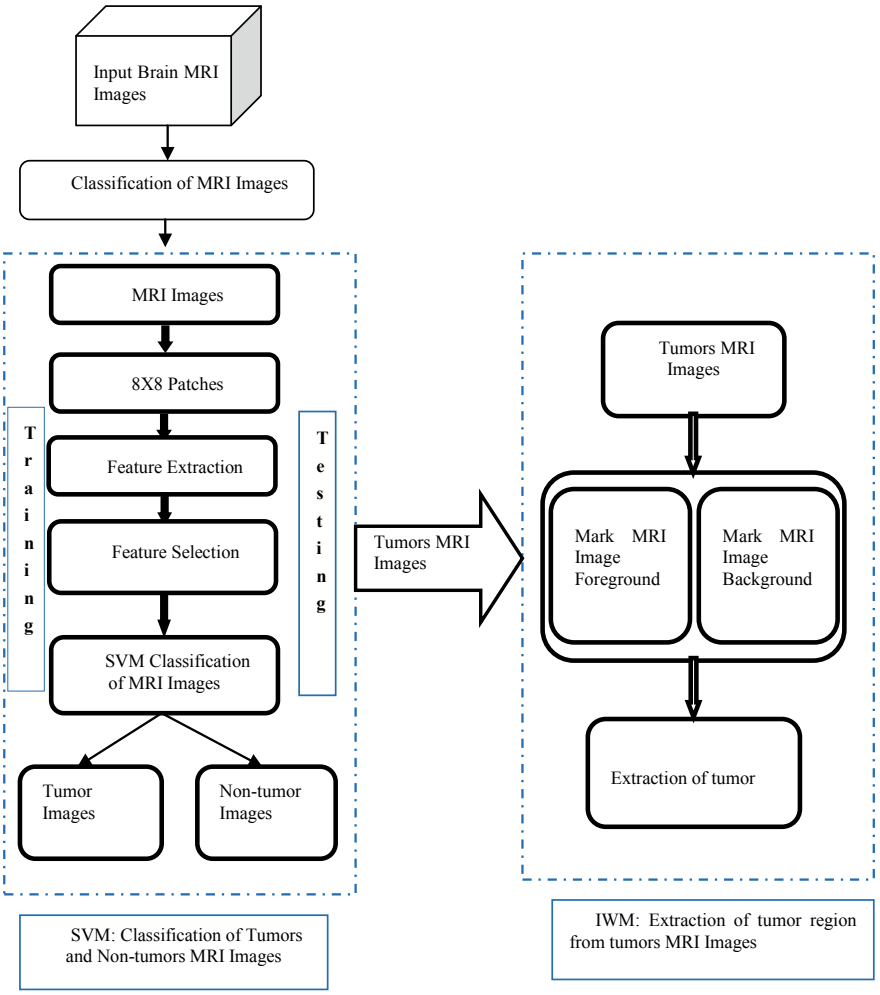


Fig. 51.2 Proposed robust framework

SVM method is used for classification of input brain MRI images into tumorous and non-tumor MRI images, and SVM [6, 9, 10, 13] is a supervised machine learning mode which is most popular for classification of MRI images.

The demonstration of SVM method for classification of input MRI images is shown in Fig. 51.3. The illustration of brain MRI images for feature selection and classification of input MRI images is shown in Fig. 51.3.

Phase-II: Extraction of Tumor Region from Tumors MRI Images with Improved Watershed Method**.**

In this research work, the improved watershed method is developed for extraction of tumor region from MRI images. The marker controlled watershed method is used for identifying or marking foreground objects and background locations for extraction of tumor in brain MRI images. The manifestation of improved watershed method for segmentation of tumor in MRI images is shown in Fig. 51.4.

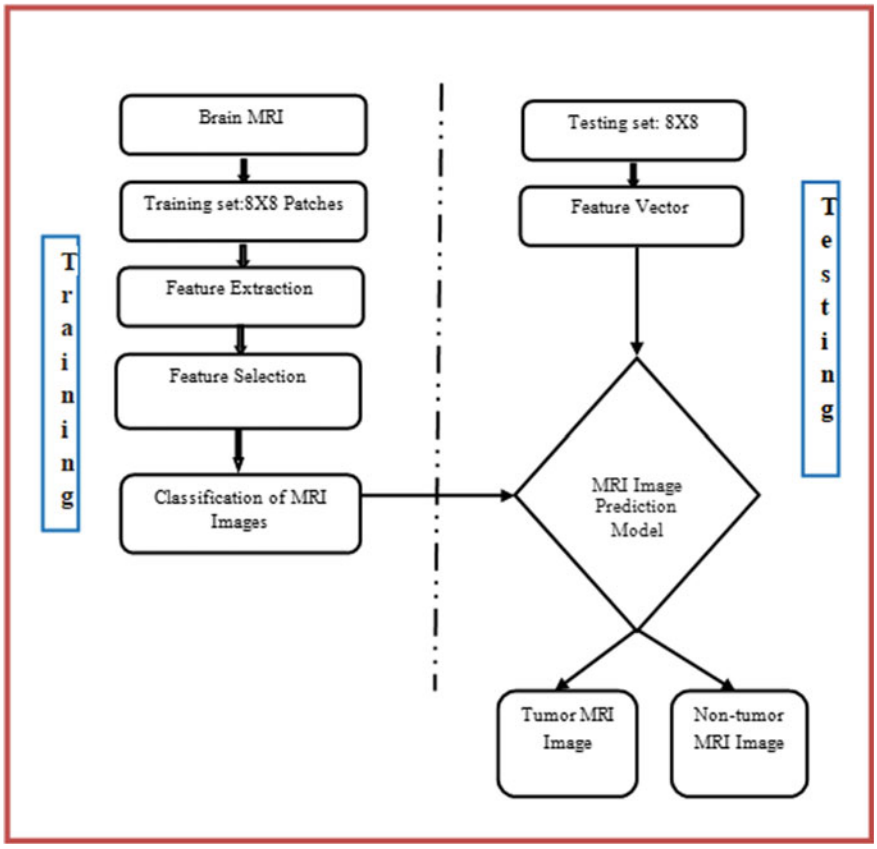


Fig. 51.3 Block diagram of SVM for MRI images classification

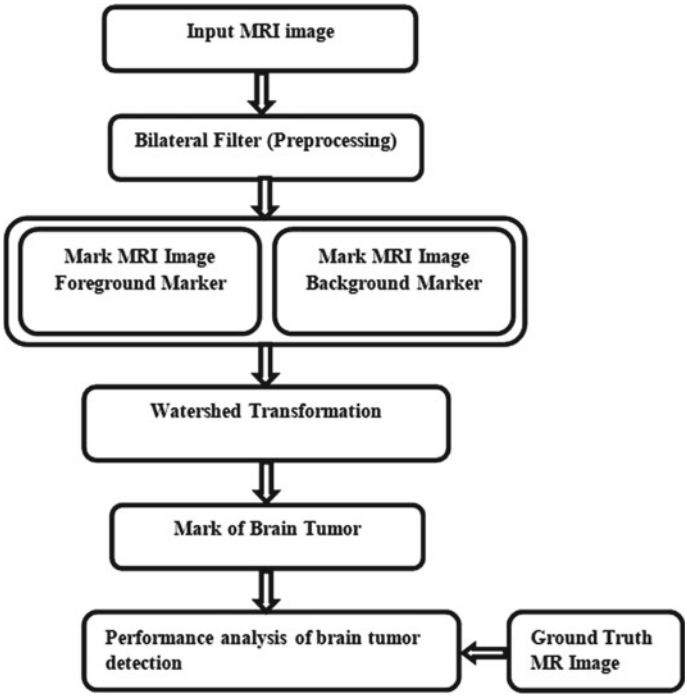


Fig. 51.4 Framework for improved watershed method for segmentation of tumor in MRI images

51.3 Results and Discussion

The research work quantitatively evaluated the performances of proposed new framework with segmentation accuracy; sensitivity and specificity [14] on benchmark brain MRI datasets consist of ground truth MRI images. The higher value of segmentation accuracy indicates that the brain tumor is segmented more accurately. The SVM method (Phase-I) and improved watershed method (Phase-II) are demonstrated on extensive publicly available brain MRI datasets. The performance of SVM Method is evaluated with segmentation accuracy on benchmark brain MRI datasets. The performance of improved watershed method is evaluated with segmentation accuracy on benchmark brain MRI datasets that comprise ground truths images.

$$\text{Segmentation Accuracy } (\%) = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} * 100$$

Table 51.1 and Fig. 51.5 present the obtained results of SVM on five brain MRI datasets for classification of tumor and non-tumor brain MRI images.

Table 51.2 and Figure 51.6 demonstrate the analysis of segmentation accuracy of improved watershed method on ten patient's brain MRI images for segmentation of

Table 51.1 Analysis of obtained classification results of SVM on five brain MRI datasets

MRI Dataset	No of MRI images	Classification of MRI Images		Classification accuracy (%)
		Tumor MRI image	Non-tumor MRI image	
Dataset-1	500	403	97	99.08
Dataset-2	500	411	89	99.16
Dataset-3	500	397	153	98.29
Dataset-4	500	409	91	99.13
Dataset-5	500	392	108	98.37

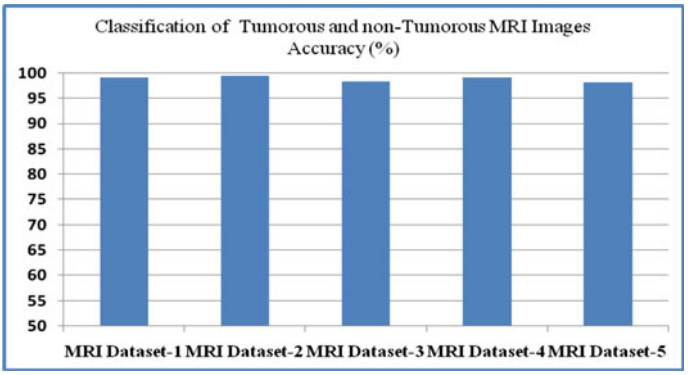


Fig. 51.5 Performance analysis of classification of tumor and non-tumor brain MRI Images with SVM

Table 51.2 Segmentation accuracy of improved watershed method on MRI images

Input MRI image	Segmentation accuracy (%)		Improvement (%)
	WM	Improved WM	
Patient-1	98.25	98.89	0.65
Patient-2	97.28	98.03	0.77
Patient-3	98.41	99.07	0.67
Patient-4	98.04	98.82	0.80
Patient-5	98.17	98.85	0.69
Patient-6	97.29	98.11	0.84
Patient-7	97.11	98.27	1.19
Patient-8	97.28	98.13	0.87
Patient-9	97.09	98.24	1.18
Patient-10	96.21	97.31	1.14

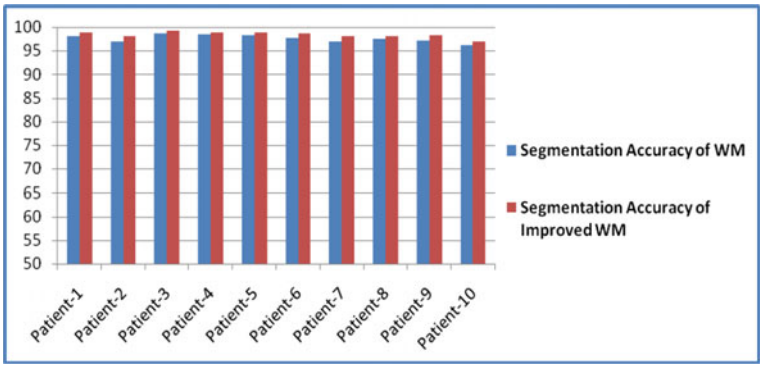


Fig. 51.6 Comparative analysis of segmentation accuracy of improved watershed method on MRI images

the brain tumor, and it is achieved improved segmentation accuracy as compared to conventional method.

51.4 Conclusion

The research work proposes a robust framework for MRI image segmentation with amalgamation of SVM and improved watershed method. The robust framework was developed to overcome the limitations of the existing MRI segmentation methods for brain MRI images. In the proposed framework, the first phase is the SVM method which is used for classification of input brain MRI datasets into tumor and non-tumor images. The second phase improved watershed method is used to automatic segmentation of tumor in brain MRI Images. The performance of proposed new framework for MRI image segmentation is evaluated on benchmark MRI dataset. The obtained segmentation results were compared with existing conventional segmentation method. The end result of the proposed robust framework presented the improvement in segmentation accuracy for brain tumor detection and segmentation.

References

1. Kalaiselvi, T., Sriramakrishnan, P.: Rapid brain tissue segmentation process by modified FCM algorithm with CUDA enabled GPU machine. Int. J. Imaging Syst. Technol. **28**(3), 163–174 (2018)
2. Kalaiselvi, T., Kumarashankar, P., Sriramakrishnan, P.: Three-phase automatic brain tumor diagnosis system using patches based updated run length region growing technique. J. Digital Imagoing Inform. Med. (2019). https://doi.org/10.1007/s10278-019-00276-2
3. Sriramakrishnan, P., Kalaiselvi, T., Rajeswaran, R.: Modified local ternary patterns technique for brain tumour segmentation and volume estimation from MRI multi-sequence scans with GPU CUDA machine. Biocybern. Biomed. Eng. **39**(2), 470–487 (2019)

4. Bhima, K., Jagan, A.: Analysis of MRI based brain tumor identification using segmentation technique. In: 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur (2016), pp. 2109–2113. https://doi.org/10.1109/ICCSP.2016.7754551
5. Kamrul Hasan, S.M., Ahmad, M.: Two step verification of brain tumor segmentation using watershed matching algorithm. Brain Inform. **5**(8) (2018). https://doi.org/10.1186/s40708-018-0086-x
6. Du, X., Li, Y., Yao, D.: A support vector machine based algorithm for magnetic resonance image segmentation. In: 2008 Fourth International Conference on Natural Computation, Jinan, pp. 49–53 (2008). https://doi.org/10.1109/ICNC.2008.400
7. Cabeen, R.P., Laidlaw, D.H.: Bilateral filtering of multiple fiber orientations in diffusion. MRI (2015). https://doi.org/10.1007/978-3-319-11182-7
8. Bhonsle, D., Chandra, V., Sinha, G.R.: Medical image denoising using bilateral filter. Graph. Signal Process. 36–43 (2012). https://doi.org/10.5815/ijigsp.2012.06.06
9. Moyano-Cuevas. J.L., et al.: 3D segmentation of MRI of the liver using support vector machine. In: Roa Romero, L. (eds.) XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013. IFMBE Proceedings, vol 41. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-00846-2_91
10. Zhang, X., Yan, L.F., Hu, Y.C., et al.: Optimizing a machine learning based glioma grading system using multi-parametric MRI histogram and texture features. Oncotarget **8**(29), 47816–47830 (2017). https://doi.org/10.18632/oncotarget.18001
11. Shih, F.Y., Cheng, S.: Automatic seeded region growing for color image segmentation. Image Vision Comput. **23**(10), 877–886 (2005). ISSN: 0262-8856
12. Bhima, K., Jagan, A.: An improved method for automatic segmentation and accurate detection of brain tumor in multimodal MRI. Int. J. Image, Graph. Signal Process. (IJIGSP) **9**(5), 1–8 (2017). https://doi.org/10.5815/ijigsp.2017.05.01
13. Blumenthal, D.T., Artzi, M., Liberman, G., Bokstein, F., Aizenstein, O., Ben Bashat, D.: Classification of into tumor and nontumor components using support vector machine. Am. J. Neuroradiol. **38**(5), 908–914. https://doi.org/10.3174/ajnr.A5127
14. Khalil, M., Ayad, H., Adib, A.: Performance evaluation of feature extraction techniques in MR-brain image classification system. Procedia Computer Sci. **127**, 218–225 (2018)

Chapter 52
Yield Estimation and Drought Monitoring Through Image Processing Using MATLAB

Shaikh Akbar Shaikh Rasul, Jadhav Swamini Narendra, and Dipti Y. Sakhare

Abstract This paper elucidates the pre-harvest yield estimation method and technique for cotton crop by image processing. This pixel-based image analysis of image processing is done using the image processing toolbox of the MATLAB 2019b. The images for abovementioned purpose are taken through camera armed drone (quadrotor). Further, there is a need for a better and transparent surveying method to assess the eligibility of a particular farm, for claiming the agricultural insurance. From the findings of proposed research, a suggestion for Agriculture Insurance Companies is made. From this research, it is concluded that yield estimation can be used for the purpose of detection of the damage and impact of the drought. The impact of drought can be assessed on individual farm level, i.e., on each acre of cotton farmland.

52.1 Introduction

This research is containing two major different parts, i.e., (i) pre-harvest yield estimation and (ii) deciding impact of the drought on individual farm level. Yield estimation will give the exact idea about the impact and situation of the drought in every individual farm. One can even decide the impact of the drought on different levels and intensities so that the more affected will get more benefits and the less affected will get less benefits. This will reduce the expenses and human efforts needed for surveying the impact of drought and also will complete the work in minimal amount of time so the farmers will not need to wait more for the government's aid. This

S. A. S. Rasul (✉) · J. S. Narendra · D. Y. Sakhare
MIT Academy of Engineering, Pune, Maharashtra, India
e-mail: shaikh.akbar@mitaoe.ac.in

J. S. Narendra
e-mail: swamini.jadhav@mitaoe.ac.in

D. Y. Sakhare
e-mail: dysakhare@etx.maepune.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_52

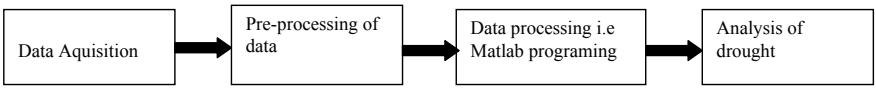


Fig. 52.1 Block diagram

method gives more precise and transparent results. Yepei et al. [1] worked on satellite remote sensing data and crop growth models, combined with the help of data assimilation technique and used SCE-UA with DSSAT to create growth model of cotton, for estimating the yield. Also, Manthan et al. [2] found a method for early prediction of drought by using standard precipitation index (SPI).

The objectives of proposed research are as follows:

1. To collect the image data of entire farm in the parts of per acre through UAV (drone armed with camera).
2. To apply image enhancement techniques as a part of preprocessing.
3. To find a novel yield estimation algorithm for estimating number of cotton bolls and thus the total estimation of yield of the farm.
4. To establish a relationship between the count of open cotton bolls and yield, at the time of harvest.
5. To use yield estimation methodology, for finding the degree of impact of drought.
6. To make an estimated yield-based suggestion, for Agriculture Insurance Companies.

52.2 Proposed Algorithm

52.2.1 Main Building Blocks

This research comprised of four distinct building blocks represented in Fig. 52.1.

52.2.2 Data Acquisition

Data is the main concept in this study. It is captured in the image form from the UAV, i.e., camera armed quadrotor. There are studies that show UAVs can be efficiently used for capturing images from the air [3]. The inclusion of drone technology for data acquisition in this study is helping us in different ways, i.e., less human efforts are required and the process becomes incredibly fast.

The size, shape and other parameters related to images that are found most suitable by performing number of experiments and referring standard available researches [4] are illustrated in Table 52.1.

Table 52.1 Specification of image capturing

S. No.	Entity	Specification
1	Camera	PX4 Optical flow sensor
2	Flight controller	Pixhawk 2.4.8
3	Secondary Processor	Raspberry Pi 3
4	Storage Unit	Micro SD
5	Image Capturing Shape	Square
6	Orientation of drone	Flat
7	Pattern of Capturing	Along the rows and columns of farm

52.2.3 Preprocessing of Data

For the betterment of the results, it is needed to make the data prepared for the processing, and this method is known as preprocessing. From the raw data captured from drone when loaded to the computer, it is preprocessed through MATLAB. Preprocessing includes different methods, i.e., (1) scaling, (2) resizing, (3) image enhancement, etc.

Scaling and resizing are crucial preprocessing techniques, especially when it comes to the tasks related to image processing and analysis. There are different levels of these methods, in simple methods, the aspect ratio is ignored. But in more advanced methods, the aspect ratio is respected.

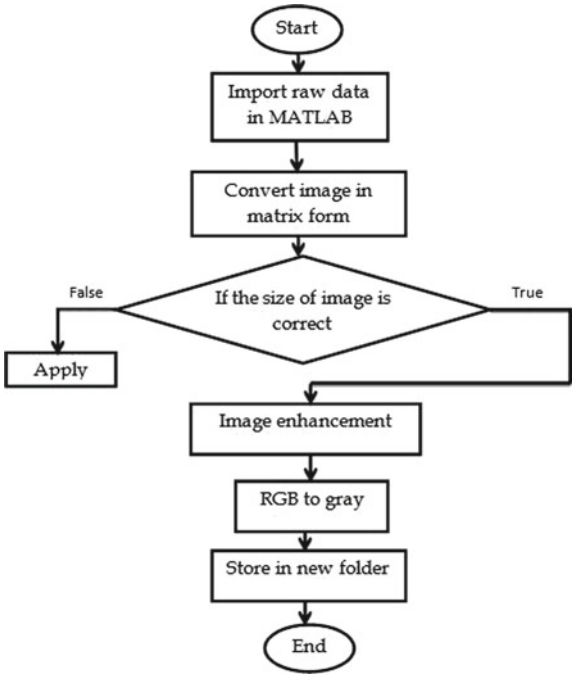
Here, scaling and resizing are done automatically as the images are taken through the same camera, and all other parameters affecting the size of the image are preset so that the captured images are of same size. But, in case if there comes any deflection in the standard size due to some errors in the hardware, it is corrected by the respected algorithm in MATLAB itself. Following flowchart depicts the process of preprocessing.

In Fig. 52.2, the overall method of preprocessing is explained through a flowchart. After importing data in MATLAB, a loop is run on the data; one image from the data is preprocessed in one iteration. Converting the image in matrix form, the next step is to check the size whether it is suitable for the algorithm or not. If it comes out to be unsuitable, then the filter is applied to correct the error. Then, some image enhancement techniques are applied for getting an enhanced image. At last, RGB image is converted into grayscale image, and all of the preprocessed images are stored in new folder.

52.2.4 Data Processing and MATLAB Programming

It contains three sub-methods which are depicted in block diagram (Fig. 52.3).

Fig. 52.2 Preprocessing



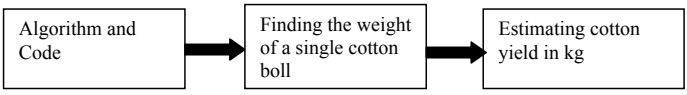


Fig. 52.3 Data processing and MATLAB programming

52.2.5 Explanation of MATLAB Code

The main intension behind this MATLAB code is to do the segmentation of the open cotton bolls, to highlight and separate them, for counting the number of these bolls. Property of the color of cotton bolls separates them from the other parts of the image [5, 6], so this property is used as the feature for the extraction of the cotton bolls from the image. According to our research, it is found that we can estimate the yield in two different forms, i.e.,

1. Number of cotton bolls per unit area
2. Amount (kg) of cotton per unit area.

But the second method cannot be directly achieved skipping the first one, i.e., the second method is the advanced version of the first method. For the first method, it is needed to extract open cotton bolls, following algorithm for the MATLAB code explains the method for counting number of cotton bolls per image.

1. Importing the image data.
2. Reading image.
3. Finding the dimensions of the image.
4. Displaying the image.
5. Performing the threshold segmentation algorithm.
6. Again displaying the threshold segmented image.
7. Thresholding the size of bolls created from the abovementioned segmentation.
8. Filtering the bolls and filling them up.
9. Using the bsxfun() for masking the image.
10. Finding centroid from the ROI, i.e., the cotton bolls.
11. Numbering and counting the bolls.

In the above algorithm, after performing the basic steps, RGB image is converted into grayscale image in preprocessing. Threshold segmentation is performed on the preprocessed image. There are two types of threshold segmentation: global segmentation and local segmentation [5, 6]. In case of global segmentation, image is divided into two regions, i.e., region of interest (ROI) and background [7]. A single threshold is used in global segmentation. For the proposed algorithm, global thresholding is used for the segmentation purpose. The grayscale intensities vary between 0 and 255, 0 representing the darkest intensities while 255 represents the brightest intensities. It is found out through experimentation that the suitable threshold value tends to 255, as the targeted part in the ROI is cotton which has the brightest intensities.

- **Experiment to find the weight of single cotton boll.**

This is needed for finding the estimation of the yield corresponding to the count of cotton bolls in the standard unit, i.e., kilograms or grams. Following steps illustrate this experiment.

1. Collected cotton in different samples, i.e., 10 bolls, 15 bolls, 20 bolls, etc.
2. Calculated the weight of these samples using digital weighing machine.
3. Then, calculated average weight of the bolls for each sample by dividing the weight of sample by the number of cotton bolls.

$$\text{Average Weight} = \frac{\text{Weight of sample}}{\text{Number of cotton bolls in the sample}} \tag{52.1}$$

4. Then, finding the average of the abovementioned average weights of each sample. This weight is considered as the final average weight.
5. In experimentation, the final average weight is found to be 4.01 g, i.e., 0.004 kg.

- **Estimating Cotton Yield in kg**

For better understanding of the yield, it is needed to represent this yield in standard units; following algorithm expresses the method to do this.

1. After getting the count of cotton bolls, the procedure for the same is explained further in Sect. 52.3.
2. Multiply the count by average weight in Sect. 2.4.2 of the single boll.
3. Get the weight of cotton present in the portion of land captured in single image.
4. Repeat the above steps for each of the captured images.
5. Sum up all the individual weights to get the overall estimation of the farm in kg.

- Analysis of Drought

There is a strong and clear relation between estimated yield and impact of drought. It can be represented in the equation form

$$\mathrm{EY} \propto \frac{1}{\mathrm{ID}} \tag{52.2}$$

where EY = estimated yield, and ID = impact of drought.

The study of the ground reality and survey with the drought affected farmers shows that there is a need to measure the impact / intensity of the drought on a scale, and this scale should include multiple discrete levels.

52.3 Experiment and Result

52.3.1 Results of Experiment Through MATLAB

One image is chosen from a dataset of images, captured from cotton farm for the purpose of experimentation. The present dataset of cotton farm is captured at the time of drought, so the yield is less, which can be observed in Fig. 52.4. Repeating



Fig. 52.4 Original image

the same process as performed below, on entire set of images for a farm, gives final yield estimation. The results are obtained using MATLAB 2019b software. There are methods, in which on-board processing is used for drone systems [8], but in the proposed research, processing is done on separate machine. The machine used for this experimentation has 8 GB of RAM and Intel i7 processor.

In Figure 52.4, this is the raw image of farm taken by camera armed drone. The size and coverage of this image are purposefully kept small for better realization of the results. Single cotton plant is covered in this image.

In Figure 52.5, here, cotton bolls in image are counted manually. So that the final results can be validated. For the sake validation, one can easily compare the number of cotton bolls counted manually and through the proposed algorithm.

Preprocessing of image is explained in Fig. 52.6. First raw image is imported in MATLAB, and then its dimensions are found and displayed that image in grayscale.



Fig. 52.5 Manually counted cotton

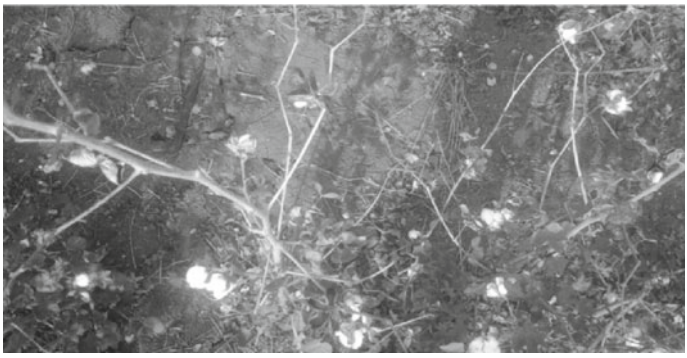


Fig. 52.6 Grayscale image

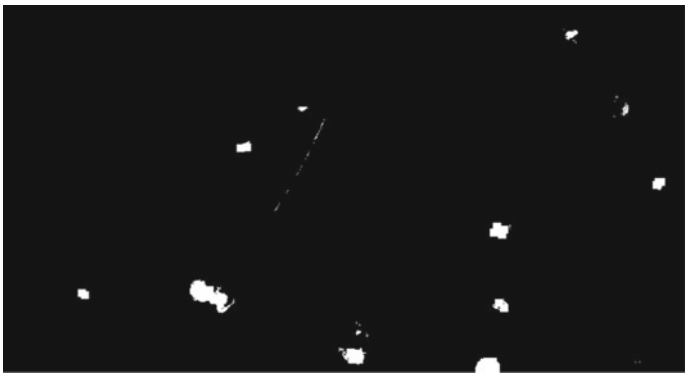


Fig. 52.7 Threshold segmented image

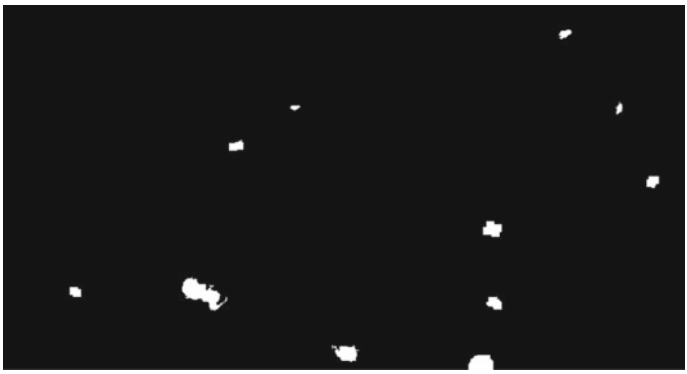


Fig. 52.8 Binarized image

In Fig. 52.7, the open cotton bolls are extracted depending upon their color properties, as they are brightest in color [7]. The intensity of open cotton bolls is around 250.

In Fig. 52.8, the color segmented image is binarized.

In Fig. 52.9, the result of finally extracted cotton bolls is shown. The cotton bolls are numbered 1, 2, 3…, and total count is displayed at corner.

52.3.2 Tables and Quantitative Analysis of Results

- After performing experiment, the average weight of the cotton bolls is found to be 4.01 g.
- Table with quantitative analysis of Sect. 53.3.1 (Table 52.2).

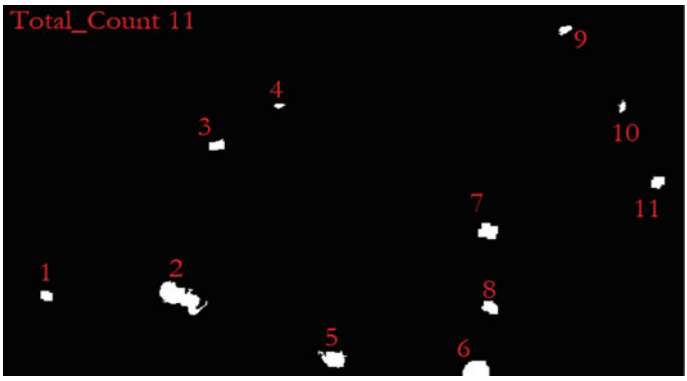


Fig. 52.9 Yield highlighted image

Table 52.2 Quantitative analysis of the experiment

Number of cotton bolls (manual)	Weight of hand-picked cotton in grams (manual)	Number of cotton bolls (automatic)	Weight by image processing in grams (automatic)
11	44	11	44.11

- Repeating the experiment in section III (A), on entire image dataset of, per acre cotton farm, gives pre-harvest final yield estimation of cotton per acre.
- The final yield estimation in Eq. 52.2 gives direct relation for estimating the degree of drought in the farm.

52.3.3 Graphical Analysis of Result

In Fig. 52.10, the outcomes of counting the number of cotton bolls through both the methods, manual and through computerized algorithm using image processing are represented using bar graph. This is the result from the experiments performed on the single image, as illustrated earlier; this image is purposefully transformed into small size for the better realization of the results.

52.3.4 Result of Another Experiment on 10 Images

Another ten random images are chosen from the dataset, these images represent 10–15 cotton plants. This data, as illustrated above, is taken from the farm which is suffering through drought situation, so the yield is less. Table 52.3 illustrates the insights of this experiment.

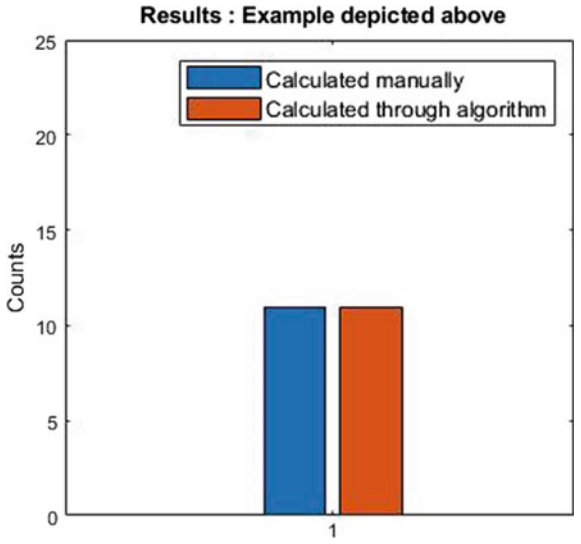


Fig. 52.10 Comparison of the results for single image

Table 52.3 Quantitative analysis of multiple images

S. No.	Number of cotton bolls (manual)	Number of cotton bolls (automatic)
1	98	95
2	105	99
3	90	102
4	97	101
5	107	94
6	85	91
7	87	88
8	90	99
9	85	89
10	85	97

Figure 52.11 is the pictorial representation of the data present in Table 52.3.

In Fig. 52.11, minor errors can be seen between both of the readings, i.e., the readings taken manually and the readings taken through proposed algorithm. In some of the samples, like in 1, 2, 3, 5, 6 and 10, it can be seen that the readings taken manually are higher than the readings taken through computerized algorithm, and this is because some of the cotton bolls which are hidden under the leaves or other things, are not visible in the images taken. Also, in other samples, like in 4, 7, 8 and 9, it can be seen that the readings taken manually are less than the readings taken through proposed algorithm, this is because some of the single cotton bolls are counted as multiple rather than single, i.e., the individual tulips of the cotton bolls separated through its own material, due to this separation, and at the time of

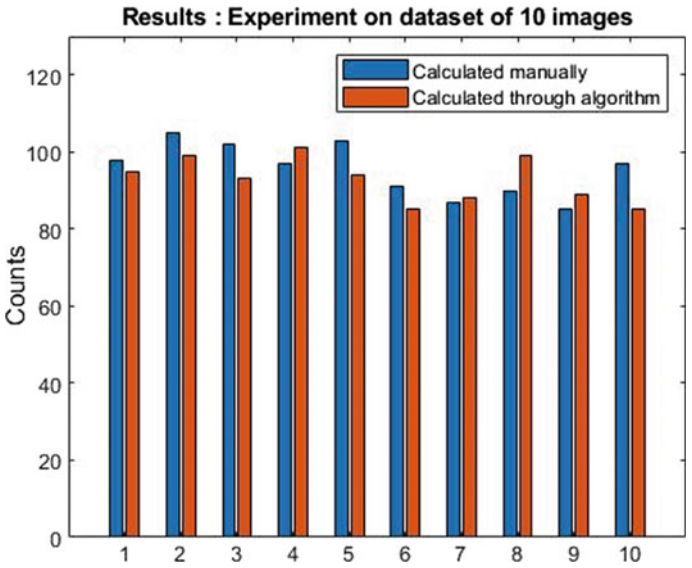


Fig. 52.11 Comparison of the results for multiple images

extraction of the cotton bolls, it is extracted as multiple cotton bolls although it is a single cotton boll in actual. For the experiment on ten samples, the total count of manual readings is 955 and the count through proposed algorithm is 928, i.e., the difference is of 27 and percentage error is about 1.43%.

52.4 Conclusion

Cotton yield can be effectively estimated through image processing, on the image dataset captured from the drone. Yield estimation has a lot of importance in the field of agriculture and drought monitoring. Drone technology reduces efforts, saves time and increases accuracy and transparency in yield estimation. Findings show that the relative yield of cotton crop can successfully estimate by considering the open cotton bolls, and still there is a space for including the closed cotton bolls and empty cotton bolls. An algorithm is designed to extract yield from individual images through image processing and clubbing them all to find yield estimation per acre. The relation between crop yield and impact of the drought is found, and a mathematical equation Eq. 52.2 is designed for the same. As per the detailed explanation given earlier, for the cause of representation, a single image is processed, similarly, the proposed novel method is implemented on another ten images, to have a big picture of method and its effectiveness. Depending upon estimated yield, impact of the drought can be expressed on different levels as low, medium and high. If the estimated yield is sufficient, then no drought condition will be announced for corresponding

farms. Therefore, the proposed research can be used to make an estimated yield-based recommendation, for Agriculture Insurance Companies. Finally, this research increases the precision and accuracy in the drought monitoring and survey. It takes down the survey level from region level to individual farm level.

References

1. Chen, Y., Mei, X., Liu, J.: Cotton growth monitoring and yield estimation based on assimilation of remote sensing data and crop growth model. In: 23rd International Conference on Geoinformatics. IEEE (2015)
2. Kansara, M., Maity, P., Malgaonkar, H., Save, A.: A novel approach for early prediction of drought. In: 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE (2020)
3. Urbanski, K.: Control of the quadcopter position using visual feedback. In: 2018 18th International Conference on Mechatronics—Mechatronika (ME). IEEE (2018)
4. Seifert, E., Seifert, S., Vogt, H., Drew, D., Van Aardt, J., Kunneke, A., Seifert, T.: Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. Remote Sens. **11**(10) (2019) (Article citation)
5. Abu Bakr Siddique, Md., Arif, R.B., Rahman Khan, M.M.: Digital image segmentation in Matlab: a brief study on OTSU's image thresholding. In: 2018 International Conference on Innovation in Engineering and Technology (ICIET). IEEE (2018)
6. Li, Y., Zhang, J., Gao, P., Jiang, L., Chen, M.: Grab cut image segmentation based on image region. In: IEEE 3rd International Conference on Image , Vision and Computing (ICIVC) (2018)
7. Gonzalez,R.C., Woods, R.E.: Digital Image Processing, pp. 711–807. Pearson, New Delhi (2017)
8. Horstrand, P., Guerra, R., Rodríguez, A., Díaz, M., López, S.: A UAV Platform based on a hyperspectral sensor for image capturing and on-board processing. IEEE Access **7** (2019)

Chapter 53
A Framework for Efficient Multilevel Polarity-Based Sentiment Analysis Using Fuzzy Logic



Shiramshetty Gouthami and Nagaratna P. Hegde

Abstract The evaluation of any product or event on social media with the opinion or emotion of peoples is known as sentiment analysis (SA). A great deal of attention has been attracted in recent years, toward both science and industry fields for a variety of uses. Machine learning and the text mining uses this most widely known application area of sentiment analysis. This paper presents a framework for efficient multilevel sentiment analysis using fuzzy logic for the classification of online test reviews polarity as strong positive, positive, negative and strong negative. This proposed model can use the fuzzy logic classifier to enhance the degree of sentiment polarity of reviews. Here, fuzzy logic classifier is used for finding the sentiment classes. This also utilizes the mechanism of imputation of missing sentiment for integrating non-opinionated sentences in generating precise results. Results show that the proposed method has a capability of extracting opinions and classify them in an effective way. The proposed method has a capability to predict the degree of sentiment polarity for the reviews on a social media. The better precision and F1-scores are obtained for an objective/subjective classification and polarity (positive/negative) classification on twitter dataset.

53.1 Introduction

In the present scenario, one of the most demanding research problems particularly in social media was sentiment analysis. Through the social media platform, peoples are able to freely convey their opinions, thoughts and observations toward various trendy dealings, subjects, products etc. by posting them on the social media. Then it has been required to analyze these posts in order to recognize which kind of sentiment that these posts convey. People emotions from a written text are recognized and

S. Gouthami
Osmania University, Hyderabad, Telangana, India

N. P. Hegde (✉)
Vasavi College of Engineering, Hyderabad, Telangana, India
e-mail: nagaratnaph@staff.vce.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_53

determined by the sentiment analysis which is also called as sentiment artificial intelligence that analyzes the views of peoples through the written posts. People around the world are allowed to freely convey their opinions on common topics through a social media platform by connecting and interacting with each other. This social media sentiment analysis can be served as a measure for determining the performance of social media in addition to use for improving customer services and marketing. Since the information in this recent years regarding any small and large incidents or disasters can be collected through social media Web sites, the effect of social media Web sites became more significant in daily life. The text data usage on the Internet nowadays increases rapidly. This enormous text data can be endeavored by the several industries to extract opinions of peoples on their products. In this context, an important source of information is the social media. Analyzing this large amount of text data manually is not possible. So it is evident that there is necessity of automatic classification.

In this case, subjective data is usually analyzed. Several social media Web sites are there where the users are enabled for contributing, modifying and rating content. There is an opportunity for the users for expressing their own views on specific topics. Social media, blogs, product review sites and forums are some of the examples of those Web sites. Twitter data has been utilized in this case. In most of the time, Web sites like Twitter contain small length of comments as status messages on social media like Twitter or article reviews on Digg. In addition, most of the sites allowed to rate the attractiveness of posts which may be related to the opinion expressed by the author. Extension for the deterministic logic gives the fuzzy logic which means that instead of a binary value, truth value has a range of 0-to-1. Conversion of white as well as black problem to a gray problem is the main goal of fuzzy logic hypothesis (Zadeh 2015). Probably the simplest method of representing human knowledge under the area of artificial intelligence is to convert it into the IF-THEN rules formatted-based natural language sentiments. These rules depend on the representations and models of natural languages and which were in turn depend on the fuzzy logic and fuzzy sets. Fuzzy rule-based classification systems are the most powerful and recognized tools for pattern recognition and classification. Since the fuzzy logic is presented in these systems, they can very efficiently handle the uncertainty and vagueness or ambiguity.

The present paper proposes a framework for efficient multilevel polarity sentiment analysis using fuzzy logic. Fuzzy logic can be used to classify the online reviews of users by considering a twitter dataset. In addition, fuzzy logic interface system generates five fuzzy-based rules to find out the sentiment for every one of the textual review tweeted on online. These reviews are classified as two classes of namely "positive" and "negative" opinions. Hence, to identify the polarities of these classified sentiments are determined fastly and efficiently using the aspect-based sentiment summarization.

53.2 Classification of Sentiment Analysis

The classification of sentiment analysis has been carried out in different ways based on the perspective point of view. Sentiment-based classification is one of those ways which is widely used as a sentiment classification technique. This sentiment classification technique is further classified into two different approaches such as the machine learning and lexicon-based approach. In addition, a third classification is also included, i.e., hybrid approach. In addition, depending on the mode of identifying and analyzing the opinions, sentiment analysis can be classified. The document level, sentence level and aspect-based levels are the three major levels of classification. Rating-based level is the last classification of the most general sentiment analysis classification. These sentiment polarities are also known as opinion polarities which were positive, negative, and neutral.

53.2.1 Document Level of Sentiment Analysis

The notions, expressions or sentiments expressed toward a component or event are the opinions. The peoples are allowed by the Web sites or forums of having a large amount of information to convey their opinion as observations and reviews. These opinions can be different from one user to another user. Positive and Negative are the just two kinds of classes observed in this analysis. One of the suitable examples is that a review on the item: “I brought a latest I Phone three days back. It is an excellent phone. It has a high-speed touch screen. It has the best voice clarity. I really love this phone". Then by using the review or star framework, required sentiments are labeled as either positive (stars rated with 4 or 5) or negative (stars rated with 1 or 2).

53.2.2 Sentence Level of Sentiment Analysis

Since this analysis can make the polarity of the sentence perfectly, it is used to provide helpful data in searches. This level of sentiment analysis can be carried out on the sentences containing opinions which may have reviews of either positive or negative. This sentence-level sentiment analysis is almost same as that of document-level sentiment analysis. However, the major difference of this sentence-level sentiment analysis is that every sentence is separately analyzed it to determine if that express a positive or negative opinion. Since the sentences can be distinguished as the objective sentences from subjective ones, this level offers more flexibility than the document level.

53.2.3 Aspect-Based Sentiment Analysis

In the consideration of a single aspect, both the document-level sentiment analysis and the sentence-level sentiment analysis are well functioned. On the other hand, peoples can able to discuss regarding the aspects in a long time that have different points of reviews or qualities. Likewise, each individual aspect would have to be unambiguous sentiments. This was generally happened when talking about social issues and in an item review. One of the convincing examples is that I am a Nokia phone of extra large. I like the look of the phone. The screen is bigger and clear. The camera is great. However, there are also some drawbacks that the battery life is insufficient and Whatsapp is not accessible. The significant data of item is hidden by requesting the positives and negatives regarding this review.

53.2.4 Comparative Sentiment Analysis

Responses of the items or brands are expressed by the users with the utilization of different expressions even it can be the same item or brand. Discovering the possibility of comparative type sentence is the main aim of this comparative sentiment analysis.

53.2.5 Sentiment Lexicon Acquisition

The process of determining expressions and opinions of reviews by utilizing the sentiments on review is called Sentimental analysis. The two types of positive and negative classes are observed in this Sentiment analysis. Let us considered a sentence "Auto X is superior to any Auto Y". This sentence does not represent which class this sentence belongs to. Similarly, the three types of systems such as manual methodology, dictionary-based approach and Corpus-based approach are used to analyze that type of sentences / texts.

- Manual Methodology: Since it is not possible to retrieve the data as a positive or negative by this method, it is completely time taking process.
- Dictionary-based approach: The SentiWordNet is used in this approach along with the Part Of Speech tagging to determine the polarity of the sentences.
- Corpus-based approach: This approach uses a domain-specific sentiment lexicon to perform the analysis.

These are the different approaches to examine the opinions of customers and also to forecast the market value of a particular organization.

53.3 Mutilevel Polarity Sentiment Analysis Using Fuzzy Logic

The proposed frame of multilevel polarity-based sentiment analysis using the fuzzy logic is shown in Fig. 53.1. The degree of sentiment polarities is enhanced by the use of fuzzy logic. The textual reviews are taken from the sentiment polarity dataset of social Web sites. This fuzzy logic can be accomplished by implementing the fuzzy interface system using a Skfuzzy package. The main aim of this proposed framework is to classify the sentiment polarities of review dataset fastly and effectively using an aspect-based sentiment summation.

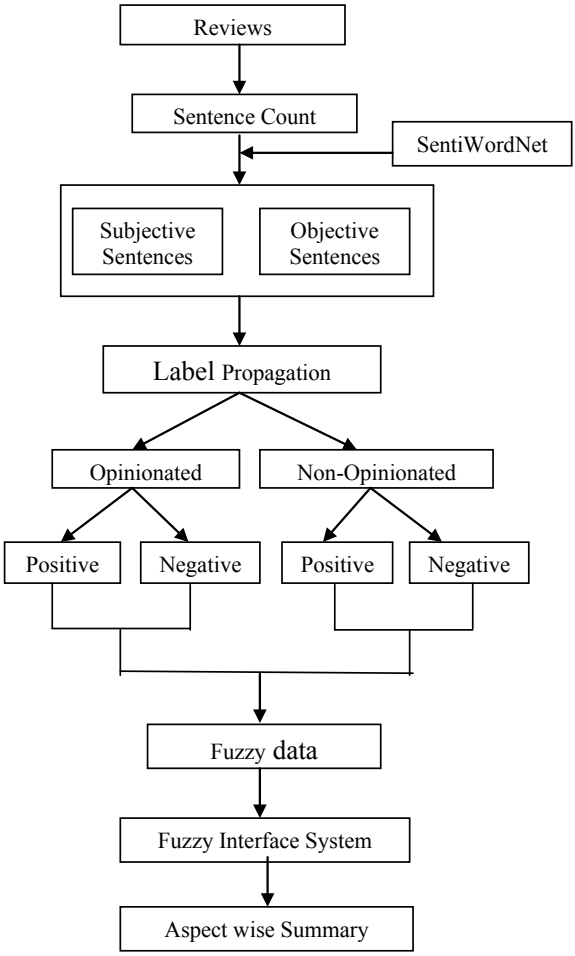


Fig. 53.1 Framework of multi level-based sentiment analysis using fuzzy logic

53.3.1 Online Reviews

The movie-related textual reviews on online that are recently collected for the opinion exchange associated with a variety of domains are used from the polarity dataset of V0.9. This dataset can be divided into training, testing and validation units. Some of the elements in the textual data such as digits, punctuations, stop words and other special symbols are removed and then processed after the completion of preprocessing. Then these text reviews from the preprocessing are divide into tokens and transformed into word vectors.

53.3.2 Sentence Separation Using SentiWordNet

The place at which a split or break occurs is called separation. A lot of sentences are present in the review dataset used in this framework. Commas, dots and semicolons are considered to separate those sentences. Moreover, in this sentence separation phase, the count of sentences in a review is also being calculated. Every one of the sentences must be split into words and word array is used to store them on behalf of future processing. The following gives the procedure of Sentence Separation and Sentence count.

```
Function Sentence_Count()
{
      Ch1='y',i=0,sen_count=0,word=0,sen[],word_array[]
      While ch1=='y' do
       Read a character ch
      if ch=='\t' || ch==' ' then
            word=word+1
      if ch=='.' or ch=='!' or ch=='?' or ch=='\n' then
             sen_count=sen_count+1
       else sen[i]= ch
      end if
      end if
      word_array[i]=ch
      i=i+1
       print 'Want to Continue?'
      read ch1
      End while
      Print sen_count
}
```

53.3.3 Sentence Label Propagation

The total numbers of sentences from a review are labeled in this process. The SentiWordNet is used in this process to label the separated sentences. If the sentences are separated as sentiment words then they are labeled with "Opinionated" or else labeled with "Non-Opinionated". The type of polarity for every opinion word from review text can be calculated by the SentiWordNet (Word). The description of sentence label propagation function is as follows,

```
Function Label_Propagation ()
Step 1: Seed SentiWordNet
Step 2: If Sentiment Word present
           Label 'Opin'
        Else
           Label 'Non-Opin'
        End if
```

Opinionated labeled sentences come under the "subjective sentences" whereas non-opinionated sentences come under the "objective sentences". The presence of aspect words in a review is determined by the use of aspect word list. The strength of a sentiment can be retrieved, if there is a presence of aspects in that, from the SentiWordNet wordlist. Objective sentences are usually accurate sentences that have no effect by their nature. However, the missing sentiment can be filled for the existed aspects by the use of imputation of missing sentiment function which can also determine the strength of the feeling from the dictionary. Following represents the imputation of missing sentiment function.

```
Function IMS ()
{
If aspect present
Impute the sentiment
Go to step 6
Else
Exit;
}
```

53.3.4 Sentiment Classifier

Each sentence can be split into tokens in which every one of tokens is labeled with polarity that representing a floating point number between 1 and −1.

Positive sentence: The review given by the user toward something with a positive or good response is considered as the Positive sentence. Example for such positive sentence can be given as,

"Most excellent movie forever" or "It was an inspirational movie!!!"

Negative sentence: The review given by the user toward opposing something with a negative response is considered as the negative sentence. Example for such negative sentence can be given as,

"Worst movie ever" or "It is waste of time"

Hence, crisp values are determined at the end of this process for these two positive and negative sentence classifiers.

53.3.5 Fuzzy Logic Interface System

The proposed sentiment analysis method using the fuzzy logic according to the SentiWordNet is described in this fuzzy logic interface system. As presented before the fuzzy logic system begins with a crisp value and after fuzzify it using different steps (fuzzification and Rules inference). Finally, a crisp value is returned in the output by the defuzzification methods (centroid, Mean/, Max…). The input (crisp value) of the fuzzy logic interface system is of two measures (positivity and negativity) calculated with SentiWordNet, and the output (crisp value) is the class of the review (positive, negative or neutral). In this case, since it is required to classify the review sentences according to two classes (positive, negative), two input variables are defined: the positive and the negative of the review sentences; and one output variable which is the class (sentiment) of the review sentences.

Every variable of fuzzy logic system is known as linguistic variable either it may related to input variables or output variables. Number of values is there in each of these linguistic variables referred as fuzzy sets or linguistic terms which can be obtained. Thus, the positive and negative are the two linguistic variables presented in the inputs of our proposed method and then low, moderate and high are the three linguistic terms for each one of those positive and negative linguistic variables. That means, probably three linguistic terms are taken by each positive and negative variables or it can be also stating that each linguistic variable is related to the three different fuzzy sets. Similarly, a linguistic variable is presented in the output which is the class of the text review and three different linguistic terms can also be taken, which are positive, negative and neutral. Then, defining the crisp values of inputs is the nest step in this fuzzy logic system which were help to start this approach after define the linguistic variables and their linguistic terms in the input and the outputs. For that, these crisp values of inputs which will have a significant part for calculating the positive and negative variables of a review use some different text preprocessing methods and SentiWordNet dictionary.

Table 53.1 Classes of fuzzy sets

Linguistic classes	Range
Very low (VL)	(0; 0; 0.3)
Low (L)	(0; 0.3; 0.5)
Medium (M)	(0.3; 0.5; 0.7)
High (H)	(0.5; 0.7; 0.9)
Very high (VH)	(0.7; 0.9; 1.0)

53.3.6 Fuzzy Sets

Fuzzy sets are capable of having elements with partial degree membership. Polarity and subjectivity crisp values of all considered text reviews are computed and converted to fuzzy sets. Let FP is a fuzzy set denoting polarity or subjectivity of review features $f_1, f_2, f_3, \cdots, f_n$ and $\mu_1, \mu_2, \mu_3, \cdots, \mu_n$ is the membership value of features. By using Zadeh's notation 11, the degree of polarity of a feature can be represented using these membership values as:

$$\text{FP} = \left\{ \frac{\mu_1}{f_1} + \frac{\mu_2}{f_2} + \frac{\mu_3}{f_3} + \cdots \frac{\mu_n}{f_n} \right\}$$

Fuzzification
The process of transforming inputs of crisp or real values to fuzzy set values with the help of partial degree membership is known as Fuzzification. The triangular membership functions are used to fuzzify the input variables of separate polarity and subjectivity values regarding the text reviews and producing the degree of polarities as an output for this model. Some of the fuzzy sets classes of both the input and the output variables separation are depicted in Table 53.1.

Defuzzication
The fuzzy rule-based interface system has a final step called defuzzification. If the equivalent fuzzy sets and related functions are given then the crisp value is an output of the defuzzification step. The crisp value is transformed back from the fuzzy values by using the centroid method as a defuzzier.

53.4 Results

The proposed framework can be studied and validated experimentally in this section. The experiment analysis can be carried out for the multipolarity sentiment analysis in terms of accuracy with the implication of fuzzy logic-based systems. Particularly, the main aim of this experiment analysis is comparing the proposed fuzzy logic method with that of other computational methods of machine learning models toward

Table 53.2 Results from polarity classification (positive/negative)

Classes	Count	Degree of polarity
Positive	700	0.821
Negative	700	0.769

the classification of sentiment analysis in terms of their classification accuracy. The polarity dataset of V0.9 regarding the movie reviews on online is used for conducting this experimental analysis. Table 53.2 depicts the number of positive and negative class instances within polarity dataset.

The sentiment classification is carried out first on the objective / subjective and then positive/negative polarities. Results from this classification are considered as the first step of proposed sentiment classification method. The above two classification approaches simply use the selected features that means only 5 features were selected for both the objective/subjective and positive/negative classifications. Then the fuzzy logic is employed at these two classification stages to improve the accuracy of sentiment classification very efficiently. The results of polarity classification are calculated by making a condition of using only the subjectively labeled reviews when reporting their results which distinguish between positive and negative classes. Then, this condition is detached in the case of the final classification approach and the objective and polarity classifications are effectively applied to all the reviews, in spite of whether they are marked as objective or subjective.

Finally, the aspect-based sentiment summarization technique is implemented after the fuzzy logic implementation in the classification process which uses an imputation of missing sentences mechanism to further improve the accuracy rate of sentiment classification in an effective manner. The proposed fuzzy logic is compared with the other two classifier models of machine learning approaches in terms of their sentiment classification accuracy shown in Fig. 53.2.

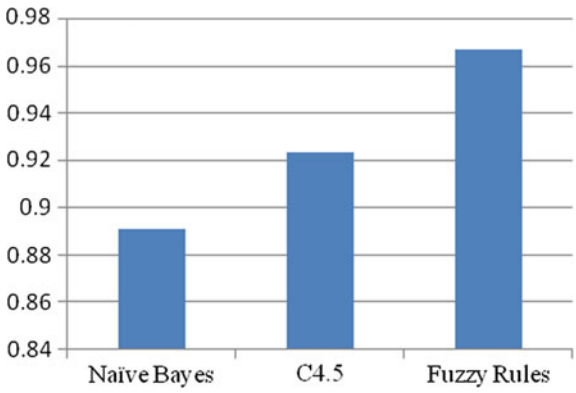


Fig. 53.2 Accuracy comparision of sentiment classification

53.5 Conclusion

A new framework is proposed in this paper for efficient multilevel sentiment analysis using fuzzy logic. The fuzzy logic used here classifies the sentiments of the user reviews. Aspect-based sentiment summarization is integrated with fuzzy logic interface system in the proposed framework for the sack of opinion mining and sentiment analysis on the online dataset. The degree of polarity for the sentiments from online reviews is then predicted in this framework. The imputation of missing sentiment mechanism that has an important part to produce accurate results is used to incorporate the non-opinionated sentences. Thus the proposed framework offers a better accuracy for the sentiment analysis on online dataset than other previous frameworks.

Chapter 54
Practical Assessment of Application Layer Protocols for Resource Constraint IoT Applications



Manasi Mishra and S. R. N. Reddy

Abstract Internet of things (IoT) systems are designed to monitor environmental parameters which are accessible either directly or indirectly through applications via Internet as a global platform. IoT devices are often equipped with limited resources requiring energy-efficient and bandwidth-efficient lightweight protocols for data transmission. Traditional protocols are not suitable for communication in IoT-based applications due to high overhead and complex architecture. The selection of effective messaging protocol highly depends on the type of application and its messaging requirements. This paper presents a broad comparison and practical assessment of the most prominent protocols such as HTTP, MQTT, CoAP, AMQP and WebSocket for IoT. A real test bed has been developed for the experimental evaluation of these protocols based on the major key performance metrics such as overhead and latency. The experimental results reveal that CoAP is the most suitable protocol for the applications having constrained resources as it introduces the least overhead. However, WebSocket is the most suitable for the time critical applications as it takes the least time for communication. The relative analysis of these protocols can help the user to choose an appropriate protocol for their application requirements and suitability.

54.1 Introduction

In today's world, the usage of IoT devices capable of sensing and communicating with each other to collect data from environment has increased exponentially. Every bit of data transferred in the network creates a carbon footprint in the environment. Even though the carbon footprint by a single IoT device is negligible, the system having huge number of devices creates a considerable carbon footprint in the environment. With huge number of IoT devices installed and sending data frequently, it is important

M. Mishra (✉) · S. R. N. Reddy
IGDTUW, New Delhi, India

S. R. N. Reddy
e-mail: srnreddy@igdtuw.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_54

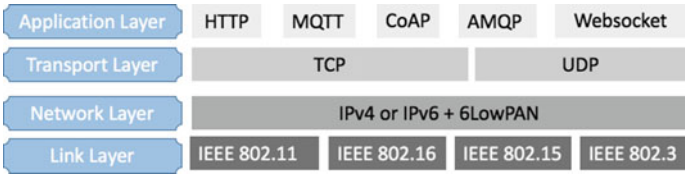


Fig. 54.1 IoT protocol stack

to choose techniques which can reduce the amount of carbon footprint by limiting the amount of data injected into the network without compromising on their functionality. The radio module is the part of IoT device which requires most of the power. Each layer of an IoT protocol stack plays a vital role in managing the communication efficiently. Figure 54.1 provides a generic protocol stack for IoT.

The protocol stack shown in Fig. 54.1, is broadly divided into four functional layers, namely (i) Link, (ii) Network, (iii) Transport and (iv) Application layer responsible for performing different subtask of communication. The characteristics such as transmission data rate, coverage area and scalability required by applications, are offered by physical layer. The IoT uses various communication technologies such as WMAN (IEEE 802.16), WPAN (IEEE 802.15) and WLAN (IEEE 802.11) for interaction among the devices at the physical layer [1]. Network layer is responsible for end-to-end data communication within the scope of the network. The traditional IPv4 addressing technology has limited address blocks which is not sufficient for IoT scenario. IPv6 standard provides a 128-bit address field, making it possible to solve the addressing issues in IoT, but it also introduces overheads which are not suitable for constrained nodes [2]. Transport layer protocol is responsible for flow control, congestion control, sequence delivery and reliability in end-to-end communication between two hosts in the network. TCP (Transmission control protocol) and UDP (User datagram protocol) are the prominent protocols at transport layer of the Internet protocol stack. Application layer is an interface between user applications and the underlying network layer providing services for application programme to ensure effective communication within the network. The selection of an effective application layer protocol is a challenging task for any developer or organization as it is highly dependent on the type of application, data sharing requirements and deployment scenario.

Some protocols have been designed for applications which need reliable and fast business transactions or operations such as AMQP (Advanced Message Queuing Protocol). Some have been designed for applications which need regular data updates in constrained network such as MQTT (Message Queuing Telemetry Transport) and CoAP (Constrained Application Protocol). Some protocols have been designed to support web applications which needs to communicate over the Internet such as RESTful client/server protocols HTTP (HyperText Teleco and CoAP). Few have been designed for the applications which require instant messaging (IM) and online discovery such as SIP and XMPP. This shows any one protocol cannot address every requirement of all possible IoT use cases. Also, it is important to investigate

the performance of these protocols in real environment to determine their best-fit scenarios [3].

The remainder of this paper is organized as follows: Section 52.2 presents the theoretical background of the five most prominent application layer protocols for IoT systems: HTTP, MQTT, CoAP, AMQP and WebSocket; Section 52.3 demonstrates an experimental test bed setup developed for practical assessment of these protocols in real environment; Section 52.4 presents the results of the experiment carried based on the key performance metrics and also presents the discussion about their suitability for certain scenario; Section 52.5 concludes the paper and suggests scope for future improvements.

54.2 IoT Application Layer Protocols

54.2.1 HTTP

HTTP is a web-based messaging protocol, standardized by IETF and W3C jointly. HTTP supports RESTful Web architecture based on request/response model. To locate the resources, HTTP uses Universal Resource Identifier (URI). Server sends data through the URI and client receives data through particular URI. HTTP is a text-based protocol and undefined header and message payloads size. Being TCP at transport layer, the communication between HTTP client and server is a connection-oriented. To make the communication secure, HTTP uses TLS/SSL [4]. It has no defined mechanism for QoS provisioning. It relies on TCP layer for packet retransmission in case of loss packets. HTTP is the most famous and accepted web messaging protocol offering several features such as chunked transfer encoding, request pipelining and persistent connections [3]. Figure 54.2 shows the architecture of HTTP protocol.

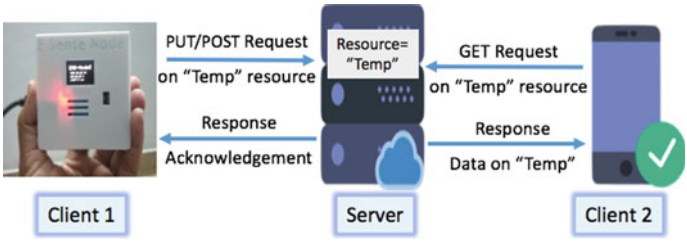


Fig. 54.2 HTTP protocol architecture

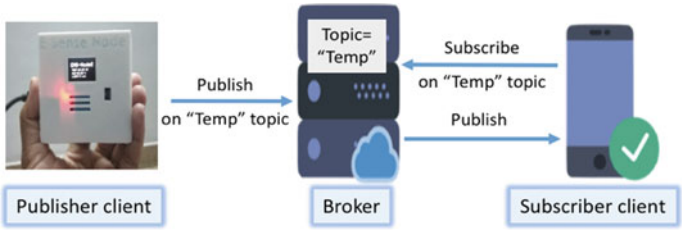


Fig. 54.3 MQTT architecture

54.2.2 MQTT

It is one of the oldest devices to device communication protocol introduced in 1999 by IBM. It is an open-source publish/subscribe model-based lightweight and simple messaging protocol designed for constrained environment. MQTT clients communicate with each other over a topic via a centralized server know as broker. Clients can be a publisher which is interested in sending messages and/or subscriber which is interested in receiving messages. The publisher publishes message to MQTT broker on an address known as topic. Every client which has already subscribed to that topic will receive the message. Having TCP at underlying layer, MQTT communication between broker and client is a connection-oriented. MQTT offers three levels of QoS for reliable delivery of messages. QoS-0, i.e. "At most once", where fire and forget mechanism is applied and message loss can occur, QoS 1, i.e. "at least once" where messages are guaranteed to arrive but redundancy can occur. and QoS 2, i.e. "exactly once" where messages are guaranteed to arrive exactly once [5, 6]. Extension of MQTT offers implementation over UDP known as MQTT-SN to make it more lightweight and suitable for resource constrained devices [7]. Figure 54.3 shows the MQTT communication architecture.

54.2.3 CoAP

CoAP is a web transfer protocol specially designed for constrained networks with constrained nodes [RFC 7252]. It is similar to the widely used HTTP for web-based applications and offers clients to use basic CRUD operations on the resource using four methods like HTTP: POST, GET, PUT, and DELETE. It uses the RESTful architecture based on request/response model making information available through resources identification by the URI [RFC3986]. CoAP builds on top of UDP, which does not provide a retransmission mechanism as provided in case of HTTP being TCP at its underlying layer. CoAP header has two bytes to set the type of message, i.e. confirmable, non-confirmable, acknowledgement and reset. To ensure the reliability through the application layer, CoAP uses exponential retransmission mechanism by setting message type as confirmable [8]. The extension of CoAP has the feature

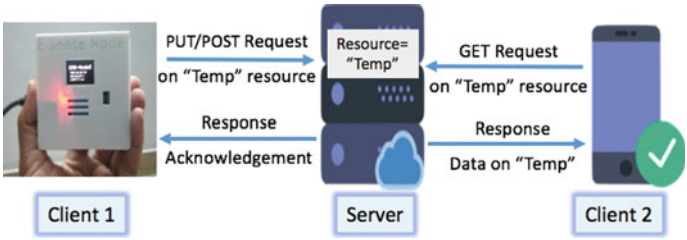


Fig. 54.4 CoAP architecture

of resource/observe to provides the functionality of publish/subscribe model while maintaining the properties of REST [8]. Figure 54.4 show the architecture of CoAP network.

54.2.4 AMQP

AMQP is a lightweight messaging protocol designed for work provisioning, reliability, interoperability and security in enterprise environment. It supports both publish/subscribe as well as request/response model for message transfer. In AMQP-based system, the publisher or consumer creates an "exchange" with a name and then broadcast them to be discovered by others. A consumer creates a "queue" and associates it to exchange immediately. Messages received by the exchange have to be matched to the queue via a process called "binding". Being TCP at the transport layer, broker and client have connection-oriented communication. AMQP is an appropriate protocol for businesses transactions as it offers wide range of services related to messaging such as topic-based publish-and-subscribe messaging, reliable queuing, flexible routing and transactions. It provides various ways to exchange route messages: directly, in fan-out form, by topic, and based on headers. AMQP offers reliability by offering three levels of QoS for delivery of messages: at most once and at least once and exactly once [9]. Figure 54.5 show the architecture of AMQP protocol.

54.2.5 WebSocket

WebSocket is a lightweight protocol offering full-duplex communication over TCP standardized by IETF [RFC 6455] to use it for real-time applications. It supports asynchronous, full-duplex and low latency communication between client and server through a single socket which was not supported in HTTP. The protocol is divided into two parts: handshake part in which client and the server establish initial communication using HTTP and data transfer part can be done using the WebSocket

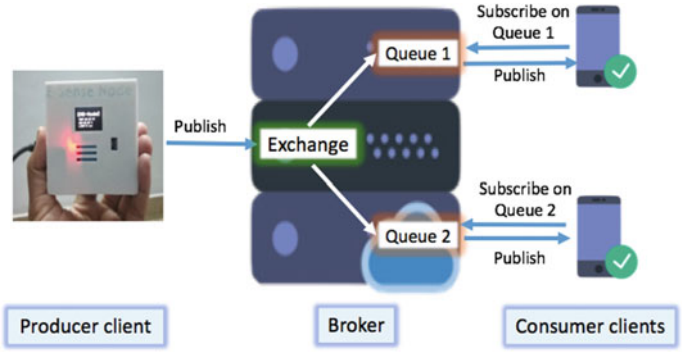


Fig. 54.5 AMQP Architecture

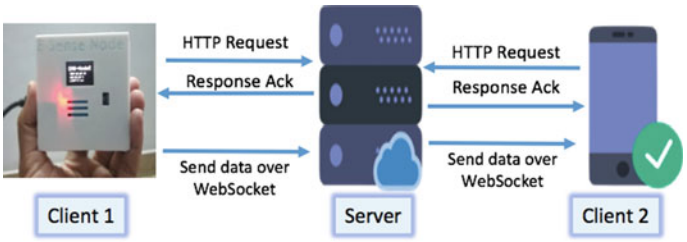


Fig. 54.6 WebSocket architecture

protocol. Messages over WebSocket are divided into one or more frames using custom binary framing format and are sent in frames, introducing only 2 bytes overhead. WebSocket customization capabilities allow compression and multiplexing through its extensions [10]. Figure 54.6 show the architecture of WebSocket protocol.

Based on the literature survey, a comparison between these protocols with respect to significant parameters has been presented in Table 54.1.

There are various broker/server implementations and client libraries available for actual use of these protocol in real-world scenarios. Most of the researchers use simulation techniques to evaluate the performance of protocol but simulation fails to mimic the real-world irregularities and hardware dependencies. Very few researches have focused on performance evaluation of these protocols on real hardware by creating a test bed. Section 52.3 presents the experimental test bed setup to evaluate the performance of these protocols on real hardware.

Table 54.1 Comparison of application layer protocols

Protocol	Standard organization	Header length	Abstraction	Transport layer	Default port	Encoding format	Message Distribution	QoS option
HTTP	IETF	Undefined	Request/response	TCP	80/443 (TLS/SSL)	Text	1: 1	No
MQTT	OASIS	2 bytes (min)	Publish/subscribe	TCP	1883/8883 (TLS/SSL)	Binary	1: Many	Yes
CoAP	IETF	4 bytes (min)	Request/response and Resource/observe	UDP	5683 (UDP)/5684 (DLTS)	Binary	1:1 and 1:Many	Yes
AMQP	ISO/IEC	8 bytes (min)	Publish/subscribe and request/response	TCP	5671 (TLS/SSL), 5672	Binary	1:Many and 1:1	Yes
WebSocket	IETF	Undefined	Request/response	TCP	80/433 (TLS)	Binary	1:1	No

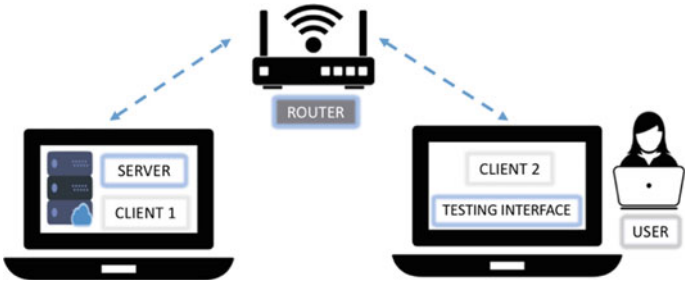


Fig. 54.7 Experimental Test Bed

54.3 Experimental Test Bed Setup

54.3.1 Scenario

The experimental test bed consists of two laptops and a Wi-Fi router as shown in Fig. 54.7. The server/broker of each protocol has been installed on one laptop and other laptop testing interface running each client application.

Client: There are two types of clients in our experimental test bed. Client-1 is acting as a receiver and client-2 is acting as a transmitter. Client 1 and client 2 can communicate via server installed on the laptop 1. The testing interface has the client module of each protocol which will be invoked based on the user choice.

Server/Broker: Server/Broker of each protocol has been installed on laptop 1.

54.3.2 Key Performance Indicator

The factors which are most significant for the choice of any protocol is considered as key performance indicator. The key performance indicators for the test are overhead and latency.

i. Latency: The difference between the time when message was received by the client 1 and the time when the message was sent by client 2 indicates the time taken for the communication.
ii. Overhead: The additional amount of data transferred for sending any message from sender to receiver is an overhead. Actual data transfer for sending a message from client 2 to client minus the message size is the indicator of overhead introduced by the protocol.

Wireshark tool has been used to capture the flow of packets across the network [11, 12]. Below parameters were recorded through the data captured from Wireshark.

1. Overhead introduced by application layer (bytes)
2. Total overhead including acknowledge and other additional packets (bytes)
3. Round trip time (mS)

54.3.3 Experiment Flow Chart

An integrated application code has been developed which can act as an easy interface for user to carry out the experiment easily just by choosing the parameters one by one through the terminal. Figure 54.8 shows the flow of experiment.

Experiment starts with starting the Wireshark application which will capture the packets flowing through the network followed by starting the test code. User can choose the protocol by which they want to send their message. Based on the protocol selection the next choices will appear on the terminal such as the name of resource

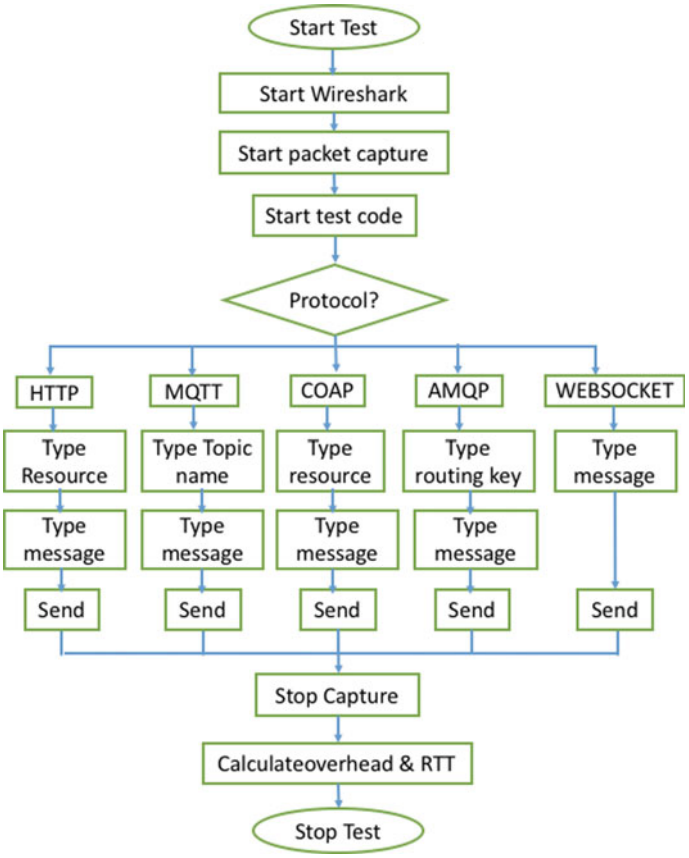


Fig. 54.8 Flow chart of experiment

Table 54.2 Experimental results

Parameter	HTTP	MQTT	COAP	AMQP	WebSocket
Overhead (bytes)	146	75	45	220	64
Total overhead including Ack (bytes)	358	143	83	377	124
Average RTT (ms)	0.91	0.52	1.61	0.43	0.32

in case of HTTP and CoAP, name of topic in case of MQTT and the name of routing key in case of AMQP. After that user will be asked to type the message they want to send. For the testing purpose, we have kept the size of resource/topic/routing key name and message constant in case of each protocol for fair evaluation. After sending the message, Wireshark packet capture can be stopped and the parameter such as overhead and RTT can be calculated.

54.4 Results and Discussion

Although the comparisons of protocols depend on the types of IoT application, resources, devices, specific conditions and requirements of the system, we have summarized the outcome of our experiment in a generic way in Table 54.2.

The results reveal that CoAP is the best protocol in terms of overhead which makes it most suitable for the applications having constrained resources. However, WebSocket is best protocol in terms of latency as its round trip time is least making it most suitable for the time critical applications. AMQP is a lightweight protocol but its support for provisioning, security, interoperability and reliability increases the overall message size and overhead [3, 4].

An important point here is that the real comparison may vary due to various other factors such as network condition, use of different server or broker implementation [12] and may differ from our results. Also, our evaluation is based on static parameters and does not consider the dynamic network conditions and overheads introduced in the retransmission of lost packets, which may also change overall overheads and amount of data transmitted thereby affecting the comparison results.

54.5 Conclusion and Future Work

This paper covers a broad comparison and practical assessment of the most prominent protocols such as HTTP, MQTT, CoAP, AMQP and WebSocket for IoT. A real test bed has been developed for the experimental evaluation of these protocols based on the major key performance metrics such as overhead and latency. The results reveal that CoAP is the best protocol in terms of overhead which makes it most suitable for the applications having constrained resources. However, WebSocket is best protocol

in terms of latency as its round trip time is least making it most suitable for the time critical applications. The relative analysis of these protocols can help the user to choose an appropriate protocol for their application requirements and suitability.

In future, we will consider all the available libraries of these protocols for the performance assessment of each protocol. We will improve the existing test bed by using emulator to mimic the actual dynamic network conditions which could help to capture the overheads incur in the retransmission of packets due to losses to get more accurate results. The performance analysis of each protocol will be done with different test cases based on the parameters such as reliability, overhead, power consumption and latency.

Acknowledgements The Authors would like to thank Mr. Rachit Thukral, Director, ETI Labs for his constant support and help where ever required.

References

1. Gazis, V., Gortz, M., Huber, M., Leonardi, A., Mathioudakis, K., Wiesmaier, A., Zeiger, F., Vasilomanolakis, E.: A survey of technologies for the internet of things. In: 2015 IEEE International Wireless Communications and Mobile Computing Conference, pp. 1090–1095 (2015)
2. Liu, C., et al.: IPv6-based architecture of community medical internet of things. IEEE Access **6**, 7897–7910 (2018)
3. Naik, N.: Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP. In: Systems Engineering Symposium (ISSE), IEEE International. IEEE (2017)
4. Luzuriaga, J.E., Perez, M., Boronat, P., Cano, J.C., Calafate, C., Manzoni, P.: Testing AMQP protocol on unstable and mobile networks. In: Internet and Distributed Computing Systems, pp. 250–260. Springer, Berlin
5. Light, R.: Mosquitto: server and client implementation of the MQTT protocol. Journal of Open Source Software **2**(13), 265 (2017)
6. [mqtt-v5.0] MQTT Version 5.0. Edited by Andrew Banks, Ed Briggs, Ken Borgendale, and Rahul Gupta. 07 Mar 2019. OASIS Standard. https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html. Latest version: https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html
7. Roy, D.G., et al.: Application-aware end-to-end delay and message loss estimation in internet of things (IoT)—MQTT-SN protocols. Future Gener. Computer Syst. **89**, 300–316 (2018)
8. Bormann, C., Castellani, A.P., Shelby, Z.: CoAP: an application protocol for billions of tiny internet nodes. IEEE Internet Comput. **16**(2), 62–67 (2012)
9. Vinoski, S.: Advanced message queuing protocol. IEEE Internet Comput. **10**(6) (2006)
10. Kayal, P., Perros, H.: A comparison of IoT application layer protocols through a smart parking implementation. In: 2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN). IEEE (2017)
11. https://www.wireshark.org/
12. Babovic, Z.B., Protic, J., Milutinovic, V.: Web performance evaluation for internet of things applications. IEEE Access **4**, 6974–6992 (2016)

Chapter 55
Collaborative-Based Movie Recommender System—A Proposed Model

Prajna Paramita Parida, Mahendra Kumar Gourisaria, Manjusha Pandey, and Siddharth Swarup Rautaray

Abstract Recommendation systems are intelligent search systems that predict the preferred information and provide suggestions to users. Major platforms such as Amazon, Netflix, Youtube improves users' experience. Recommender engines operate on background, reading users' behavior, and suggesting items that are most likely to engage with. Approaches like content filtering lead to ambiguous content or item cold-start issues. To tackle the demerits of sparsity of data and cold-start, the paper focuses on a system with a memory-based collaborative approach, employed with a similarity analysis of users' perspectives using Pearson correlation. In addition to this, a systematic study and a proposed model are presented along with the memory-based algorithm. The study shows that the Pearson correlation approach can achieve better research objectives. The user preferences have been collected from various microblogging platforms to analyze the present trends and responses of users. Promising results are produced by the experiments done over the public database.

55.1 Introduction

Many information-based companies utilize recommendation systems on a wider scale such as Twitter, Google, Netflix, LinkedIn, etc. This area of recommender originated in the mid of the 1990s, along with the invention of Tapestry, the first Recommendation System. Movie recommendation plays a major role in the mobile environment. The idea of being valuable came from the Netflix prize contest in 2009 for 1 million dollars to improve the engine, which was the biggest leaps of

P. P. Parida (✉) · M. K. Gourisaria · M. Pandey · S. S. Rautaray
Kalinga Institute of Industrial Technology (Deemed to be University), Bhubaneshwar, Odisha, India
e-mail: 1964003@kiit.ac.in

M. Pandey
e-mail: manjushafcs@kiit.ac.in

S. S. Rautaray
e-mail: siddharthfcs@kiit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_55

the techniques used. It carries out the aggregation of reviews and users' preferences to help them for better movies and to maximize profit by minimizing risks. This system provides selective information taking prior user habits and history as inputs. However, it requires both timeliness and accuracy. Researchers studied the application of machine learning algorithms as the field of recommendation developed. Since the late 1950s, artificial intelligence got emerged and then machine learning came into existence. An increase in ML algorithms is experienced such as clustering, k-nearest neighbor, Bayes network, etc. The parameters like timeliness and accuracy of the recommender system will be improved [1].

Various information of users is collected from tweets, social media sites, etc. A web-based portal is used for micro-blogging and social networking tasks that provide enormous real-time user preferences [2].

Figure 55.1 is the illustration of two approaches of content-based and collaborative-based. Content-based depends upon users' history including director, actors, description, genres, use tags for analyzing attributes or features. Whereas collaborative studies information or behaviors from other users. The main idea is to leverage the behavior pattern and predict suggestions. Broadly there are two collaborative-based methods: User-to-User and Item-to-Item approaches shown in Fig. 55.2. These methods provide intuition that people might choose items similar to other items already being chosen and also the items that are chosen by other people with similar tastes.

Item-to-Item-based method has been commonly used in areas such as LinkedIn (2014), YouTube (2010), Amazon (2003), etc. Item-to-Item-based system surfaces similar items from key values data. They are more interpretable as the scores of key values data are too small because of the selected threshold. In the user-to-user approach given in Fig. 55.2, identification of users is done based on the most similar interaction profile by evaluating distances between users. This user-centered system finds the similarity of a specific user of interest with others. They consider a similar interaction on the same items or choices for estimating closeness based on a similar

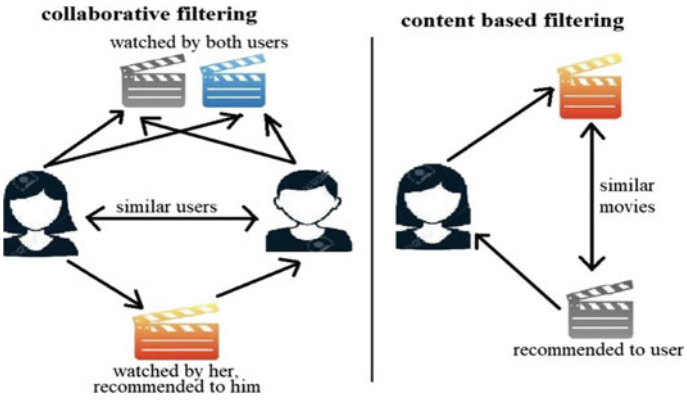


Fig. 55.1 Collaborative and content-based approaches for movie recommendation

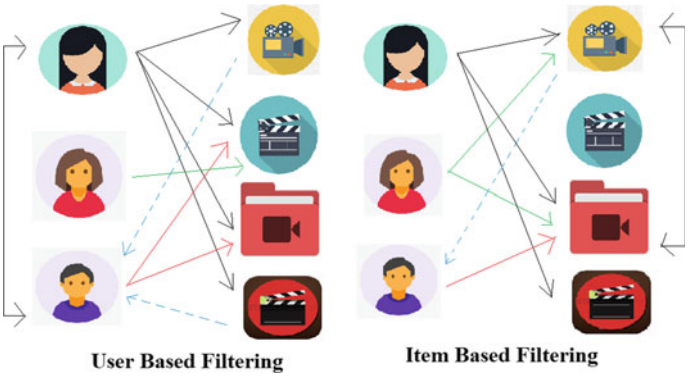


Fig. 55.2 User-Based and Item Based Collaborative Recommender system

rating or time hovering. For example, the nearest neighbor algorithm refers to those items which are not been interacted with. Whereas the item-centered works for item similarities checking whether most users have interacted with both the items in a similar way. They are less personalized and more biased but more robust compared to user-centered techniques. In this paper, item-to-item collaborative is framed as it provides better outcomes.

The remaining paper is demonstrated as given: Sect. 55.2 describes a survey over the literature part, Sect. 55.3 presents the proposed model, and Sect. 55.4 illustrates the set of selected features. The implementation and algorithm that obtains the working of the proposed model are illustrated in Sect. 55.5 along with the results in Sect. 55.6 and then in the last section, the conclusion is drawn.

55.2 Literature Survey

Different approaches have been proposed for suggesting recommendations. In the field of business applications, Guo et al. [3] suggested a novel recommender system that compares pre-purchase ratings with post-purchase ratings so that preferences can be more accurate. The system procures its instantaneous data from the virtual environment of 3D products. It mainly takes the latent choices of the user. It deploys a new mechanism of collecting pre-purchase emotions from electroencephalogram (EEG) signals that are set on a headset. In this way, these pre and post ratings can boost the overall working of the recommender. Pearson correlation computes the similarities for nearest neighbors and provides both ratings benefiting e-business. Chiru et al. [4] presented a movie recommender for upgrading a personalized type of recommendation. Various factors like user profile, history preferences, and movie scores are analyzed. A psychological test such as a questionnaire is performed and a similar profile brings out similar choices for genres. Four psychological types such as phlegmatic, choleric, melancholic, and sanguine and their intensity has been

estimated. Genres and characteristics are categorized according to these traits. So, in this case, the absence of ratings can be easily dealt with, because the association of user profile with movie probabilities could be considered. But, it fails for accurate value due to the traits evaluation.

Jenq-Haur et al. [5] presented a system to estimate the exact rating and more accuracy of classification by using sentiment lexicons. Opinion aggregation and regression techniques are employed as mostly review contains more opinion rather than ratings. This evaluates review polarity to find a compound score for recommending movie preferences. Most people share informal comments or words as opinions. Frangidis et al. [6] comparatively studied reviewers and movie scripts using Gini-index or SVM approach for increased accurate ratings. The correlation was reflected in the predictions which were carried by sentiment analysis using vector semantic and meta-features. Tools were applied like VADER (Valence Award dictionary) for sentimental analysis and NRC for emotion analysis. A mixture of machine learning algorithms is used such as Multinomial Naive Bayes (MNB), Logistic regression, SVM, Multilayer perceptron. Among these, Multinomial Naive Bayes provided promising results. Deldjoo et al. [7] studied a combination of different features of meta-data and audio–video resources mainly to tackle cold start issues. An extraction from videos constitutes a movie genome is done following the canonical correlation method.

This model focuses on items that are combined with interaction taking item content descriptors and making it quite bias-free. Thus works for collaborative enriched content with high coverage as a multimodal recommender system. Soleymani et al. [8] analyzed the movie scenes and its ranking, provided by the emotion of users and features of videos. In this system, physiological signals such as galvanic skin resistance, respiration pattern, electromyograms, body temperature, blood pressure were recorded and stored. The physiological signals were applied to categorize and rank the contents of the video to obtain optimal performance with the best features. Wang et al. [9] proposed a recommendation model that works on social content. This system evaluates the user-content matrix and finds out the relevance between video and users for re-sharing suggestions. Based on the user-video matrix, the social content area is constructed to verify relevance. Following the content delivery network, the dynamics of video sharing are high and access patterns are analyzed for matrix modification. Nidhi et al. [10] experimented on real-time data like tweets for generating a Naive Bayes based classifier. The system is applied on a small train set, providing better accuracy. To improve individual techniques, hybrid systems are composed of combinations of multiple recommendation approaches.

Melville et al. [11] developed a new hybrid approach named content-boosted CF, that utilized content-based features in a collaborative model. They overcame the first rater and sparsity problem by building a pseudo matrix of ratings. Then, the model is compared with pure content, collaborative, and naive hybrid systems. Yang et al. [12] referred ratings from page count which are collected from users' data. The measure of likeness of documents by users is demonstrated by the count of read ones. This concept focuses on the cold start issue in CF and utilizes the mean AP correlation method over log information. Zhang et al. [13] demonstrated a framework by the interaction of user recommender which takes user input, suggests a particular no of

items to the user, records choices of the user, and uses a random and kNN algorithm. In this paper, a movie recommendation framework has been suggested by associating the collaborative method with users' behavior eliminating content-based issues.

55.3 Proposed Model

The generalized data used in this model has been taken from two sources: the movie database (TMDB) and MovieLens, which fetches demographic information for users. The dataset from MovieLens is publicly available containing 26 million ratings for 45,000 movies given by 270,000 users. A small dataset is also available containing 100,000 ratings for 9000 movies given by 700 users. One file with TMDB Id of every movie is listed in the MovieLens dataset. After scraping of data, the stringified features are converted into a suitable form for parsing. Figure 55.3 is obtained showing the frequencies of ratings and how it depreciates with the increasing number of users. This presents the data sparsity phenomenon, also said to be a long plot distribution as shown in below figure.

There are few ratings by a small no of users and the majority rest are much less known with little or no user interactions. This leads to sparse ratings which pose problems and become less predictable for most users. This scarcity works opposite to abundance state. It is observed that the curve behavior has a stiff fall from the 'head' to the longtail. So, filtering is done multiple times to remove unnecessary attributes and to reduce the dimension of the dataset and sparsity as well.

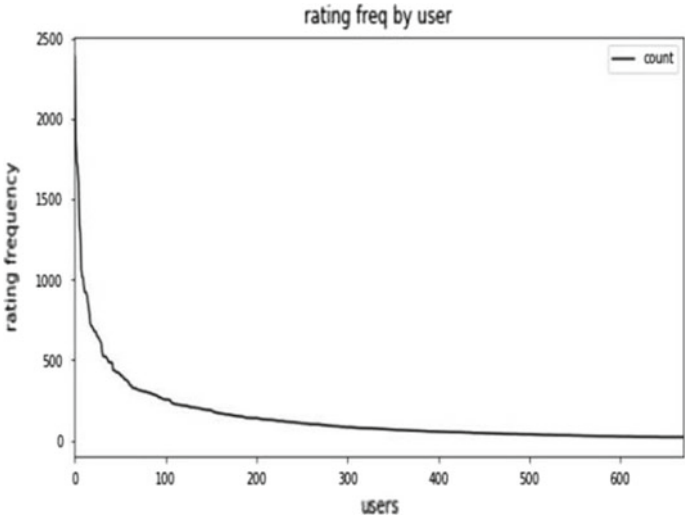


Fig. 55.3 Ratings frequencies by users

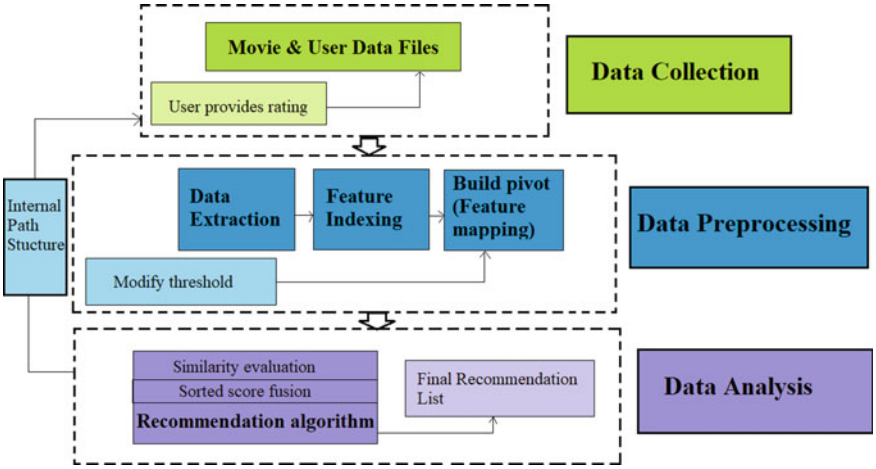


Fig. 55.4 Workflow diagram for movie recommendation system

To improve personalized recommendations, this model of collaborative approach has been proposed considering user's perspectives. Figure 55.4 depicts the procedure of movie recommendations in a stepwise manner involving data collection, data preprocessing, and data analysis.

The data collection section depicts the retrieval of data from the databases having user-rated movie data and users' profile data. These are the dataset that has to be procured and saved for updating. Users keep on adding their ratings or reviews which are also updated. Data preprocessing includes the feature analysis by filtering and indexing to get pivot data. It filters noises and unnecessary attributes by removing hashtags, repetitive words, emojis, replacing NaN values with zeroes, etc. Normalization provides matrices to estimate similarity for score calculation by maintaining threshold. The final task is data analysis that calculates scores from Pearson correlation and ranking is done. Pearson is an advanced form of cosine similarity calculation that automatically standardize the ratings data. As there are large variations of users' ratings with too much harshness or leniency. Then the similar estimated ratings are analyzed according to the user's request to finalize the corresponding suggestion list. There also lies an internal path structure that connects the first section with the last section to update the ratings data. This calculates the compound score and evaluation of similarity between relevant movies are processed. In this procedure, a collaborative-based approach is applied to produce a preliminary suggested list. Recently, for an exponential increase in the online data, this system works effectively for decision-making. This proposed method brings convenience and suitability compared to content-based.

Table 55.1 Feature set for rating estimation

Features	Description
Id	Movie ID by TMDb
rating	Ratings given by specific user
userId	user's identity
movieId	Movie id for TMDb
tagline	Specific front detail of the movie
title	Official name of the movie
genres	category of movie
timestamp	Duration when ratings are given
vote_average	Average reviews of movie
vote_count	No. of votes given by users
imdbId, tmdbId	Movie id given on respective platform

55.4 Feature Selection

As the data are being collected, it requires integration, cleaning, transformation, segmentation, stemming, and removal of noise for the pre-processing phase [14]. Thus to make the data in a form that it would be easily parsed through the machine and to obtain accuracy, data are transformed into clean data. The processed data discover the pattern and becomes understandable to interpret. The following features being listed in Table 55.1, are mainly taken into consideration for extraction purposes in the recommendation model. Many attributes or features have to be filtered from the main dataset.

55.5 Implementation

To find users' preferences, the collaborative approach is considered combining the content features of metadata with the users' interactions. This calculates the similarity function of two items at a time in relevance to the target user and then estimates the user-item matrix by Pearson-correlation. The Pearson correlation method gets the measure of similarity of one entity with other nearby entities. Neighboring vectorized estimations are obtained and these scores are ranked by mapping functionalities. Similar interests of different users are considered. Thus, taking movie-user similarity parameters, feature vectors are accessed, and based on rating, a final combined rating matrix is constructed. We intend to present the below Pearson-based algorithm in systematic procedure. This could be further extended for different comparisons with other approaches by accuracy measures on various data and correlation coefficients.

Pearson-Based Algorithm for the Proposed Recommender System

Inputs:

Dataset containing CSV files
Metadata of movie rating
Normalized similarity score.

Mains:

1. Build a pivot table using the ratings matrix.
2. Set threshold to reduce dimensions of sparse data.
3. Calculate Pearson similarity correlation to find angular distance between items.

Mathematically, given as:

$$r = \frac{N\sum xy - \left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[N\sum x^2 - \left(\sum x\right)^2\right]\left[N\sum y^2 - \left(\sum y\right)^2\right]}} \tag{55.1}$$

where

N no. of pairs of scores
x, y scores or entities like items.

1. Standardize ratings with the similarity scores to estimate average weighted rating.
2. Sort and rank the mapped matrix to give a suggestion list.

Output:

Preliminary recommendation list
Predicted ratings.

55.6 Simulation Results

The ratings data are filtered from the metadata and the equivalent lists of users who have contributed their reviews are recorded. As the data is sparse, the review count has to be wrangled with precision to get a pivot table. Then similarity vectors are retrieved to get mapped titles by Pearson similarity index. Following dynamic ratings, the scaling is done and most similar high-rated choices are listed at the top. Figure 55.5 gives the suggestion list as per the varied choice of low to high average weighted ratings. Here, the movie request (say 'Iron Man') at the top is provided with a high rating of 5, which in return gets the corresponding resultant list in descending order of estimated ratings. Similarly, to further cross-check, when the same movie

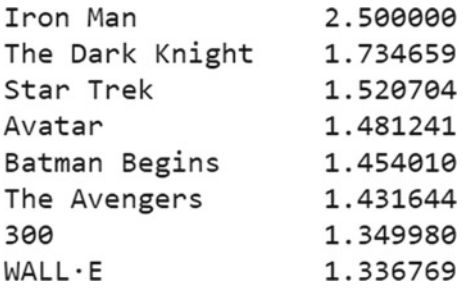


Fig. 55.5 Recommendations as per the user choice of a movie (provided high rate)

request is provided with a low rating of 1, then the resultant list gives the opposite of the old list, as shown in Fig. 55.6. Thus, it provides an increasing ordered list of ratings by putting the movie with a rating as 1 at the lowest level.

So, to evaluate the correctness of the recommender system, accuracy measures like precision and recall are applied and values are obtained, shown in the Table 55.2. These estimate the effciency of the system providing correct or incorrect decisions. Mathematically given by Eqs. 55.2 and 55.3:

$$\text{Precision} = \frac{\text{Relevant recommended items}}{\text{All recommended items}} \tag{55.2}$$

$$\text{Recall} = \frac{\text{Relevant recommended items}}{\text{All recommended items}} \tag{55.3}$$



Fig. 55.6 Recommendations as per the user choice of a movie (provided low rate)

Table 55.2 Empirical analysis by precision and recall metrics

Splits	Precision	Recall
Fold-1	0.59563	0.21788
Fold-2	0.62539	0.23241
Fold-3	0.59660	0.22530
Fold-4	0.59955	0.22778
Fold-5	0.61686	0.23555

55.7 Conclusion and Future Work

Thus, the movie recommendation framework being proposed searches preferred movies from real-time responses. The objective is to analyze the development of the recommender system that employs a machine learning algorithm and assistance to data mining areas. The main purpose is to provide a better knowledge management system and to improve the accuracy and timeliness compared to previous models. It filters unusable data by delivering personalized ideas. So, useful movie suggestions will be available to the users readily. Recommendation procedure is a comprehensive work that involves distinct users along with movies, that are collected from reviews on social media platforms. Types of correlation coefficients could be found out by considering sentiment ratings and IMDb movie ratings. This proposed model captures the change in interests of the users and works best for two unrelated factors. It will illustrate the efficiency and precision of enormous real-time data [15]. Also, it is beneficial as it takes variations of sentiments as positive, negative, and neutral sentiments by rating measures. The proposed framework could be improved with more datasets and more recent movies and we expect more extraction of information from different social media platforms to improve this system. However, the study of requirements, design, and maintenance suggests more research opportunities to be investigated and could be more explored to work for a dynamic environment.

References

1. Florea, A.C., Anvik, J., Andonie, R.: Spark-based cluster implementation of a bug report assignment recommender system. In: International Conference on Artificial Intelligence and Soft Computing (ICAISC), LNCS, vol. 10246 (2017)
2. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Analyzing temporal dynamics in Twitter profiles for personalized recommendations in the social web. In: WebSci '11: Proceedings of 3rd International Conference on Web Science, Koblenz, Germany, Koblenz, Germany (2011)
3. Guo, G., Elgendi, M.: A new recommender system for 3D e-commerce: an EEG based approach. J. Adv. Manage. Sci. **1**(1), 61–65 (2013)
4. Chiru, C.G., Dinu, V.N., Preda, C., Macri, M.: Movie recommender system using the user's psychological profile. In: IEEE International Conference on ICCP (2015)
5. Wabg, J.H., Liu, T.W.: Improving sentiment rating of movie review comments for recommendation. In: IEEE International Conference on Consumer Electronics Taiwan (ICCE- TW) (2017)
6. Frangidis, P., Georgiou, K., Papadopoulos, S.: Sentiment analysis on movie scripts and reviews. In: IFIP International Conference on Artificial Intelligence Application and Innovations (AIAI 2020), pp. 430–438 (2020)
7. Deldjoo, Y., Dacrema, M.F., Constantin, M.G.: Movie genome: alleviating new item cold start in movie recommendation. User Model. User-Adapted Interact. **29**(2), 291–343 (2019)
8. Soleymani, M., Chanel, G., Kierkels, J.J., Pun, T.: Affective ranking of movie scenes using physiological signals and content analysis. In: Second Workshop on Multimedia Semantics, pp. 32–39 (2008)
9. Wang, Z., Sun, L., Zhu, W., Yang, S., Li, H., Wu, D.: Joint social and content recommendation for user-generated videos in Online social network. IEEE Trans. Multimedia **15**(3), 698–709 (2013)

10. Nidhi, R.H., Annappa. B.: Twitter-User recommender system using tweets: a content-based approach. In: International Conference on Computational Intelligence in Data Science (ICCIDS) (2017)
11. Melville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. **23**, 187–192 (2002)
12. Wei, B., Wu, J., Yang, C., Zhang, Y., Zhang, L.: Cares: a ranking-oriented cadal recommender system. In: *Ninth Joint Conference on Digital Libraries*, pp. 203–212. ACM (2009)
13. Zhang, H.R., Min, F., He, X., Xu, Y.Y.: A hybrid recommender system based on user-recommender interaction. Math. Probl. Eng. (2015)
14. Katarya, R., Verma, O.P.: An effective collaborative movie recommender system with cuckoo search. Egypt. Inform. J. **18**(2), 105–112 (2017)
15. Rautaray, S.S., Agrawal, A.: A real-time hand tracking system for interactive applications. Int. J. Comput. Appl. **18**(6), 28–33 (2011)

Chapter 56
Classification of Bruised Apple Using Ultrasound Technology and SVM Classifier



Gopinath Bej, Tamal Dey, Abhra Pal, Sabyasachi Majumdar, Rishin Banerjee, Devdulal Ghosh, Vamshi Krishna Palakurthi, Amitava Akuli, and Nabarun Bhattacharyya

Abstract Bruising on apples is mainly caused due to mechanical damage during harvesting and postharvest journey to the supermarket. Apple's bruise damage reduces the quality of apple as well as market price. This paper proposes an ultrasonic signal analysis technique to classify bruised and unbruised apples. An ultrasound signal with 100 kHz center frequency has been applied on the apple surface and the response signal has been recorded in reflectance mode. Principal Component Analysis (PCA) has been applied to the acquired response signals to reduce the dimensionality of the dataset. The first fifteen PCA components that contain more than 99% information have been used as the working dataset for classification. Then, the Support Vector Machine (SVM) classifier has been applied to this reduced dataset to classify bruised apple. SVM is trained with the training dataset that consists of data

G. Bej (✉) · T. Dey · A. Pal · S. Majumdar · R. Banerjee · D. Ghosh · V. K. Palakurthi · A. Akuli · N. Bhattacharyya
Centre for Development of Advanced Computing, (C-DAC), Plot E-2/1, Block GP, Sector V, Bidhannagar, Kolkata 700091, India
e-mail: gopinath.bej@cdac.in

T. Dey
e-mail: tamal.dey@cdac.in

A. Pal
e-mail: abhra.pal@cdac.in

S. Majumdar
e-mail: sabyasachi.majumdar@cdac.in

D. Ghosh
e-mail: devdulal.ghosh@cdac.in

V. K. Palakurthi
e-mail: vamshi.palakurthi@cdac.in

A. Akuli
e-mail: amitava.akuli@cdac.in

N. Bhattacharyya
e-mail: nabarun.bhattacharya@cdac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_56

from both bruised and unbruised apples. The model developed after SVM training, is used to classify the test dataset. The classification result shows promising accuracy in the tune of 98%. This ultrasonic technique can endow with a new paradigm for automatic non-destructive identification of bruise-damaged apples.

56.1 Introduction

Apple is a delicious fruit rich in vitamins, minerals, and antioxidants. It is mostly consumed fresh by consumers. Fresh fruits are susceptible to bruising which is caused due to the mechanical damage of fruit flesh. The journey of an apple from the orchard to the retail shelves involves several intermediate stages like harvesting, sorting, packing, storing, and shipping [1–3]. During those intermediate stages of the journey to the supermarket, apple countenance impacts, and compression several times that cause bruise damage on the apple. Impact bruising occurs when an apple falls freely onto a hard surface and compression bruising occurs when apples are pressurized with force [4, 5]. Bruise-damaged fruits are susceptible to the blue mold decay and invisible bruise inside the skin can be susceptible to fungus infection. Consequently, a major quantity of total produced apple is wasted due to this bruise damage and imparts a substantial economic loss to the industry [5–7]. The bruising defect is a major concern to the industry and it is considered as a major quality defect. These reduced quality apples are not well accepted among the consumers and unbruised apples command the best price due to better quality [8, 9].

The ultrasonic technique is a trend to evaluate the quality of agro produces in a non-destructive manner [10, 11]. In this paper, a novel ultrasound-based methodology has been proposed for the classification of bruised apple. The ultrasound response signal is recorded in reflectance mode from the apple surface by applying an ultrasound signal of 100 kHz frequency. For every sample under the experiment, an ultrasound response signal has been acquired with the same device and the same setup. This acquired response signal from an apple surface is a one-dimensional array of length 2000. Instead of extracting features from the response signal, all 2000 data points are used as a feature vector. Principle component analysis (PCA) has been applied to the acquired signal to visualize the effectiveness to form the clusters with bruise-damaged and unbruised apples. Also, PCA reduces the dimensionality of the dataset, by considering the first few principal components which together explain most of the information. The reduced dataset has been used as the feature set for the development of the SVM classifier. After the supervised training of the SVM, the developed model has been tested against a known test dataset. The classification result shows remarkable accuracy. The ultrasonic analysis technique is a rapid and reliable technique to distinguish between bruised and unbruised apples.



Fig. 56.1 Few of the collected samples

56.2 Methodology

56.2.1 Sample Collection

For this experiment, a total of 154 numbers of apples were collected from the local market which contains both unbruised and bruised apples. Samples were marked with numbers from 1 to 154 for further traceability. During the experiment, the ultrasonic response signal of 154 apples has been acquired sequentially. Figure 56.1 is showing the images of a few unbruised and bruised apples used during experimentation.

56.2.2 Ultrasonic Setup

The experimental setup comprises three components, i.e., an ultrasound pulser/receiver (Model: USB-UT350, Make: US Ultratek), an ultrasound transceiver probe (Model: X1020, Make: Olympus Scientific Solutions Americas Corp.), and a computer. This ultrasound pulser/receiver is commercially available and it can generate a wide range of ultrasound signals by varying the configurable parameters. It can also receive the ultrasound signal sensed in the transducer probes. The ultrasound probe used in this experiment has a central frequency of 100 kHz which practically fires the ultrasound. It can also sense the ultrasound response signal when acts as a receiver. The ultrasound pulser/receiver is connected to the computer using the USB interface and the ultrasound probe is connected to the ultrasound pulser/receiver using a BNC cable. Figure 56.2 embodies the laboratory setup with the ultrasound

Fig. 56.2 Laboratory setup



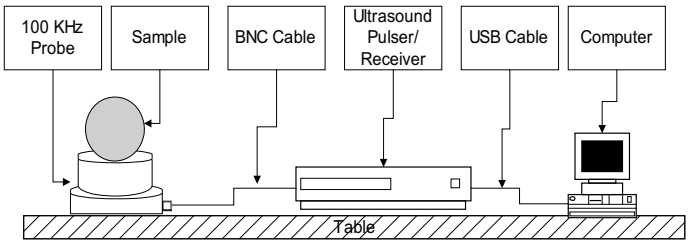


Fig. 56.3 Block diagram of setup

probe, pulser/receiver, and an apple sample placed on the probe. Figure 56.3 shows the graphical representation of the block diagram of the ultrasound setup. This same setup has been used for acquiring the data from all 154 samples.

56.2.3 Signal Acquisition

First, the sample under consideration is placed above the probe and the ultrasonic gel is used as a coupling medium to remove any air gap between the apple surface and sensor probe. The gel reduces the attenuation of ultrasonic sound waves. One software has been developed using MS VC++© to interface with the ultrasound pulser/receiver. The software is configurable to adjust the parameters for the ultrasound pulser/receiver such as buffer size, sampling rate, low pass filter, high pass filter, rectifier, transducer mode (transmittance/reflectance), center frequency, pulse-width, and pulse-voltage. as per the requirement. In this experiment, 100 kHz tone burst frequency is triggered via the probe in reflectance mode. The probe acts as a transceiver which also senses the response signal and it is acquired in the software. The ultrasound signal is subjected to the selected location on the apple surface. A recorded ultrasound response signal during experimentation is shown in Fig. 56.4.

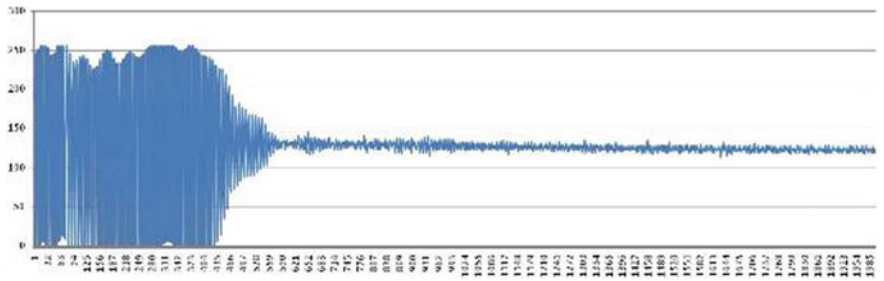


Fig. 56.4 Acquired ultrasound signal

56.2.4 Manual Analysis

After the experiment with ultrasound, each sample is cut with a knife and examined manually for bruise damages. During manual analysis, among 154 apples, 93 are classified as unbruised, whereas, 61 apples are considered as bruise-damaged apples. During the development of the classification model, the Class-I apples are considered as unbruised apples, while, bruised apples are marked as Class-II.

56.3 Data Analysis

The acquired response signal is a one-dimensional array of length 2000. For 154 samples, the acquired dataset is a matrix of dimension 154 × 2000. Instead of mining the signal specific features by applying digital signal processing and analysis technology, all 2000 data points of a response signal are considered as 2000 features of the apple. With this hypothesis of 2000 feature vectors, data analysis has been performed on the 154 × 2000 dataset for the classification of bruised apples. Initially, Principle Component Analysis (PCA) has been applied on the 154 × 2000 dataset to find out the pattern of the data points. Graphical representation of PCA components in a 2-D or 3-D plot helps to visualize the data clustering pattern. The clustering nature of data points in these plots edifies to decide on applying a classifier algorithm. Also, the dimension of the dataset is reduced based on the scores of principal components. This reduced dataset is used for the development of the classifier model. Finally, the Support Vector Machine (SVM) classifier has been used as the decision support model for the classification of bruised apples. The result represented in the Results and Discussion section proves the practical applicability of the hypothesis. A brief depiction of PCA and SVM is described below.

Principle Component Analysis (PCA) [12–15]: It is a well-established dimensionality reduction technique, used to identify correlations and patterns in a dataset. It reduces the dimensionality of a dataset to a significantly lower dimension minimizing the loss of information. Applying the underlying mathematics, PCA creates new uncorrelated variables that successively maximize variance. Each uncorrelated variables are principal components where first principal component is the direction in space along which projections have the largest variance. The steps involved in transforming the initial dataset to a reduced dataset using PCA are described below:

i. Normalize the whole dataset
ii. Compute the covariance matrix of the normalized dataset
iii. Compute the eigenvectors and eigenvalues of the covariance matrix
iv. Sort the eigenvalues in descending order and accordingly order the eigenvectors.
v. The eigenvector with the highest eigenvalue is the most significant and thus forms the first principal component.

vi. Reducing the dimensions of the dataset by selecting the first few principal components which explain the dataset with a significant score.

Support Vector Machine (SVM) [16–19]: It is a supervised machine learning algorithm used for classification problems and regression analysis. It determines the optimal hyperplane between classes which separates the data points into classes. Data points from each class that are closer to the hyperplane are the support vectors. Using these support vectors, the margin of the classifier is maximized to find the optimal hyperplane. Removing any support vector data point from the dataset influences the position of the hyperplane. During the training of the SVM algorithm, the optimal hyperplane is determined which is the decision boundary for classes. While testing the new data point, the class of the data is determined using the optimal hyperplane determined during the training of SVM.

56.4 Results and Discussion

Principle Component Analysis (PCA) has been applied to visualize the data pattern in 2-D (using the first two principal components, explains: 91.7%) and 3-D (using the first three principal components, explains: 96.2%) plots. The generated plots using MATLAB are shown in Figs. 56.5 and 56.6.

Both plots, 2-D, and 3-D represent the formation of two distinguishable clusters. Moreover, the clusters are linearly classifiable in 2-D as well as 3-D planes. The explained score with the first fifteen principal components is more than 99%. These components are considered as the reduced dataset in an uncorrelated vector space. Thus, PCA helps to reduce the dimension of the whole dataset from the 154 × 2000 to 154 × 15. Table 56.1 describes the PCA component explained scores with the first 15 PCs.

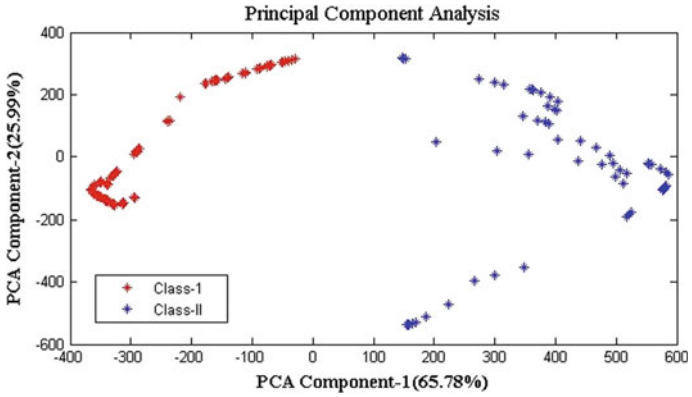


Fig. 56.5 2-D PCA plot with the first two components

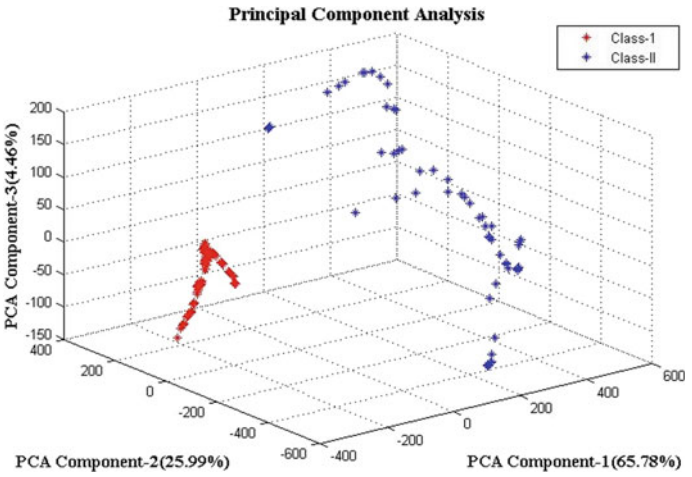


Fig. 56.6 3-D PCA plot with the first three components

The support vector machine (SVM) classifier algorithm is used with the reduced dataset for distinguishing the bruised apples. As SVM is a supervised learning classifier algorithm, it is required to be trained before the prediction of unknown samples. First, the dataset is divided for training and testing of the SVM as explained in Table 56.2. The training dataset ensembles 108 samples (70%), whereas, the rest 46 samples (30%) are used during testing of the model. Out of 108 training data samples, 65 data samples are selected randomly and belong to Class-I. Rest 43 data samples belong to Class-II. During testing, out of 46 samples, 28 samples are selected randomly from Class-I, and the rest, 18 samples are considered as Class-II. The SVM model is developed using the MATLAB® software tool.

The selection of training and testing samples is shuffled randomly for 10 times to obtain the average accuracy. The accuracy of testing using the SVM classifier with ten randomly selected datasets is shown in Fig. 56.7. The performance statistics related to the testing accuracy is shown in Table 56.3. The average accuracy to classify the bruised apples using the SVM classifier is calculated as 98.71%.

56.5 Conclusion and Future Work

The use of an SVM classifier for the classification of bruise-damaged apples using ultrasound technology shows a promising result. Moreover, the process of determining the bruised apples using ultrasound technology is non-destructive. This experiment shows a pathway for determining the intrinsic qualities of apple further in detail. In the future, data from more samples will be analyzed using this technology. Also, lower ultrasound frequency, i.e., 40, 60, 80 kHz will be explored for this study which may further be used to explore more intrinsic quality parameters of apple in

Table 56.1 PCA component explained scores (first fifteen PCs)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
65.781	25.987	4.460	1.086	0.604	0.252	0.187	0.125	0.111	0.094	0.082	0.075	0.065	0.063	0.055

Table 56.2 Dataset distribution

	Available dataset	Training dataset	Testing dataset
Class-I (unbruised)	93	65	28
Class-II (bruised)	61	43	18
Total	154	108	46

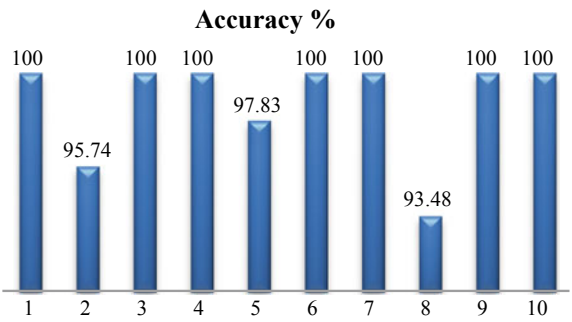


Fig. 56.7 Result of SVM classifier using random training and test dataset (repeated 10 times)

Table 56.3 Statistics of testing accuracy

S. No.	Parameter	Accuracy %
1	Minimum	93.48
2	Maximum	100.00
3	Average	98.71
4	Standard deviation	2.20
5	Median	100.00
6	Mode	100.00

a non-destructive manner. Further, this technology can be equipped with a conveyor for the online sorting of bruised apples.

Acknowledgements The authors are grateful to ICAR, New Delhi for funding this project and providing valuable guidance toward the development of non-destructive methodologies.

References

1. Hussein, Z., Fawole, O.A., Opara, U.L.: Harvest and postharvest factors affecting bruise damage of fresh fruits. Horticultural Plant J. **6**(1):1–13., doi: https://doi.org/https://doi.org/10.1016/j.hpj.2019.07.006

2. Fu, H., Karkee, M., He, L., Duan, J., Li, J., Zhang, Q.: Bruise patterns of fresh market apples caused by fruit-to-fruit impact. Agronomy **10**, 59. https://doi.org/10.3390/agronomy10010059
3. Khan, A.A., Vincent, J.F.V.: Bruising and spliting of apple fruit under ini-axial compression and the role of skin in preventing damage. J. Texture Stud. **22**(3) (1991). https://doi.org/10.1111/j.1745-4603.1991.tb00018.x
4. Mitsuhashi-Gonzalez, K., Pitts, M.J., Fellman, J.K., Curry, E.A., Clary, C.D.: Bruising profile of fresh apples associated with tissue type and structure. Appl. Eng. Agric. **26**(3), 509–517 (2010). ISSN 0883-8542
5. Lewis, R., Yoxall, A., Marshall, M.B., Canty, L.A.: Characterising pressure and bruising in apple fruit. Wear **264**(1–2), 37–46 (2007). https://doi.org/10.1016/j.wear.2007.01.038
6. Altisent, M.R.: Damage Mechanisms in the Handling of Fruits
7. Rostampour, V., ModaresMotlagh, A.: Evaluation of the bruising susceptibility of apple in transport conditions. Bulg. J. Agric. Sci. **24**(5), 902–908 (2018)
8. Dobrzanski Jr., B., Rabcewicz, J., Rybczynski, R.: Handling of apple, transport techniques and efficiency vibration, damage and bruising texture, firmness and quality. In: B. Dobrzański Institute of Agrophysics of Polish Academy of Sciences. ISBN 83-89969-55-6
9. Komarnicki, P., Stopa, R., Szyjewicz, D., Kuta, Ł., Klimza, T.: Influence of contact surface type on the mechanical damages of apples under impact loads. Food Bioprocess. Technol. **10**, 1479–1494 (2017). https://doi.org/10.1007/s11947-017-1918-z
10. Mizrach, A.: Ultrasonic technology for quality evaluation of fresh fruit and vegetables in pre- and postharvest processes. Postharvest Biol. Technol. **48**(3), 315–330 (2008). https://doi.org/10.1016/j.postharvbio.2007.10.018
11. Vasighi-Shojae, H., Gholami-Parashkouhi, M., Mohammadzamani, D., Soheili, A.: Ultrasonic based determination of apple quality as a nondestructive technology. Sens. Bio-Sensing Res. **21**, pp 22–26 (2018). https://doi.org/10.1016/j.sbsr.2018.09.002
12. Jolliffe, I.T., Jorge, C.: Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A.37420150202 (2016), https://doi.org/10.1098/rsta.2015.0202
13. Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. Chemometrics Intell. Lab. Syst. **2**(1–3), 37–52 (1987). https://doi.org/10.1016/0169-7439(87)80084-9
14. Joliffe, I.T., Morgan, B.J.T.: Principal component analysis and exploratory factor analysis. Statistical Methods Med. Res. **1**(1), 69–95 (1992). https://doi.org/10.1177/096228029200100105
15. Vlachos, M.: Dimensionality reduction. In: Sammut, C., Webb, G.I. (eds.) Encyclopedia of Machine Learning. Springer, Boston, (2011). https://doi.org/10.1007/978-0-387-30164-8_216
16. Yue, S., Li, P., Hao, P.: SVM classification: its contents and challenges. Appl. Math. Chin. Univ. **18**, 332–342 (2003). https://doi.org/10.1007/s11766-003-0059-5
17. Shmilovici, A.: Support vector machines. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook. Springer, Boston, (2005). https://doi.org/10.1007/0-387-25465-X_12
18. Birajdar, G.K., Mankar, V.H.: Subsampling-based blind image forgery detection using support vector machine and artificial neural network classifiers. Arab. J. Sci. Eng. **43**, 555–568 (2018). https://doi.org/10.1007/s13369-017-2671-3
19. Cervantes, J., García-Lamont, F., Rodríguez, L., Lopez-Chau, A.: A comprehensive survey on support vector machine classification: applications, challenges and trends. Neurocomputing (2020). https://doi.org/10.1016/j.neucom.2019.10.118.doi:10.1016/j.neucom.2019.10.118

Chapter 57
An Artificial Neural Networks Feature Extraction Approach to Predict Nephrolithiasis (Kidney Stones) Based on KUB Ultrasound Imaging



Gollapalli Sumana, Giri Aparna, and Gade Anitha Mary

Abstract Around 10% of the population will experience nephrolithiasis or renal calculi or kidney stones at their life time. Nephrolithiasis is the main cause for the haematuria, micturition and sever pain in the abdomen and during urination. More number of nephrologists are required to meet the rapid increase in the population. This paper deals with a supportive diagnostic system to the physician or nephrologist by applying neural network techniques. Generally, patients are supposed to undergo ultrasound scan test to identify the position, size and number of renal calculi in the renal system. Image processing techniques like preprocessing, segmentation and feature extraction were applied on the ultrasound scan images to extract the GLCM features. These extracted features were supplied as inputs to construct neural network and train the system by using feed-forward back-propagation algorithm. The proposed system is constructed with 22 input nodes, ten hidden nodes and one output node and thereby trained until it gets target or output. Then this trained neural network is allowed automatically to work on emerging new samples. Around 50 real-time images were collected from the patients, in which 60% of them were used for training and 40% for testing. The supportive software tool used for this proposed system is MATLAB 8.5.

57.1 Introduction

In medical diagnosis, ultrasound is protective and is painless, which produces images of internal organs of the body by capturing the sound waves. It is also identified as

G. Sumana (✉)
Sri Padmavati Mahila Visvavidyalayam, Tirupati, Andhra Pradesh, India
e-mail: sumana@spmvv.ac.in

G. Aparna
University College of Engineering, Hyderabad, Telangana, India

G. Anitha Mary
Loyola College, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_57

ultrasound scanning mainly and also as sonography by its working principle. In this process of image capturing, a small transducer (test) and ultrasound gel that is applied the skin directly are used. The sound waves of high recurrence are transmitted from the test through the gel into the body. The transducer captures the sounds that reflect, and the sound waves that are produced are used to generate an image. Since this process of examination does not make use of any ionizing radiation, therefore, it is a harmless process as there is no radiation penetration into the patient body. The ultrasound scanners constitute of a machine and electronic devices like computer, monitor and a transducer which are used to perform the scanning process. The transducer is a tiny and handy gadget which is like a microphone connected to the scanner by a cord. Different transducers are used based on the type of examination during a specific diagnosis. The transducer conveys imperceptible, high recurrence sound waves into the body and tunes later for the returning echoes from the tissues in the body.

57.1.1 Methodology

In the process of ultrasound scan test, the individual has to lie down straight facing up position on the examination table. This table is flexible in rotation position and can be tilted so as to move on either side for quality image capturing. After the initial examination of the patient, the radiologist or sonographer will apply semiliquid water-based gel smoothly to the area of the body being contemplated. The gel will help the transducer reach the body, and the air bubbles are removed which are between the transducer and the skin that can obstruct the sound waves from penetrating into individual under examination. The transducer is to be moved to and fro on the surface of the body over the range of interest until the coveted images are captured. The patient under test does not experience any uneasiness from pressure as the transducer is pressed in the area to be inspected, notwithstanding, if a territory of delicacy is being examined. The patient may feel the pressure or minor torment from the transducer. Figure 57.1 shows the pictorial representation of the medical setup of the ultrasound scanning. It has an ultrasound machine, image acquisition device, display, tracing and analysis system. The other apparatuses are a transducer with magnetic field receiver, non-ferrous bed and magnetic field generator. The advantages of the ultrasound when compared to other examination processes are ultrasound filtering which is non-invasive meaning no needless or infusions so it is harmless and not painful. It is most accessible, simple to use and cost effective when compared to other imaging techniques. It is an extremely safe method as it is not accompanied with any ionizing radiation. It produces clear images of soft tissues which are not that clear in X-ray images. It is a preferred imaging methodology for the diagnosis and for monitoring of pregnant women and fetus who should be prone to radiation much. It provides a real-time imaging, making it a handy apparatus for directing insignificantly advanced methods, such as needle biopsies and fluid aspiration.

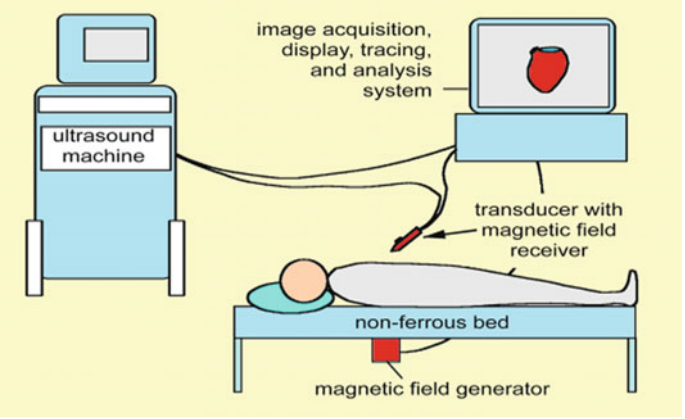


Fig. 57.1 Pictorial representation of the medical diagnosis of the ultrasound scanning

57.2 Automatic Extraction of KUB Ultrasound Image Features Using Digital Image Processing Techniques

57.2.1 Introduction

Ultrasound is utilized to assess an individual local kidneys (the ones you were conceived with) and, in addition, transplanted kidneys [1]. It can measure the size and appearance of the kidneys and recognize tumors, congenital anomalies, swelling and blockage of urine stream. A more up-to-date system called Color Doppler is utilized to survey clusters, narrowing, pseudo-aneurysms in the arteries and veins of the first and transplanted kidneys. This is the minimum invasive of all the procedures. It is compact and, most importantly, does not require radiation. As far as anyone is concerned up until this point, there have been no reports of reactions from the strategy itself. With a specific end goal to perform analysis of nephrolithiasis in view of ultrasound, initially, it is essential to extract the features from the ultrasound image [2]. Predicting stone in the renal system is not a simple task since it is unrealistic to specifically distinguish the kidney stone in the ultrasound in light of the fact that as it is a rigid mass like different organs of the kidney [3]. So to differentiate the kidney stone with other organs, ultrasound image feature extraction is especially needed [4, 5]. Consequently in the present review, by utilizing the digital image processing strategies, the ultrasound image features are extracted automatically [6].

57.2.2 Data

The data utilized for the present analysis are the ultrasound images, which are gathered from the patients of different hospitals. In order to process, these ultrasound images are stored in JPG format.

57.2.3 Procedure

Since the acquired ultrasound image may contain noise, this should be preprocessed to eliminate noise from the image. Then this de-noised image is resized through equal m rows and n columns for quick computation. In this way, first the input image is changed to grayscale image from RGB, then this image is used to do further process because the time taken by a grayscale image for processing is minimized when compared with the other images. Then this grayscale image is preprocessed to eliminate noise from the image, and then this preprocessed image is further for the extraction of the GLCM features.

57.2.4 Extraction of Ultrasound Image Features

57.2.4.1 Preprocessing

Initially, the input image which is taken is resized to quick computation. At that point, the image is changed over from RGB to gray in order to increase extraction method on the basis of the gray-level co-occurrence matrix (GLCM), or gray-level spatial dependency matrix. The vital idea of GLCM is that the texture information contained in an image is characterized by the adjacency connections that the gray tones in an image need to each other. These features are extremely robust to orientations effects of the processing speed. Then the gray image is preprocessed to remove the noise content in the image. Digital images are liable to have different varieties of noises. These noises result errors in the image procurement which in turn fails to get the pixel values that do not reflect the originality input image. Noisy images are difficult to find the abnormality. In this manner, the initial step is preprocessing, which is concluded by padding the noisy image. New pixels which exists around the boundaries of an image will be detected by the image padding. These edges act like marginal or boundary lines by utilizing progressed filtering techniques. This is an iteration process in which the aggregated de-noised pixels are generated to update the accumulated weights.

Once the image is de-noised, then the following stride is the masking so as to discover the reference. At that point of the region of interest (ROI) is taken, it is converted to binary image. At that point, the properties of ROI are extracted. These

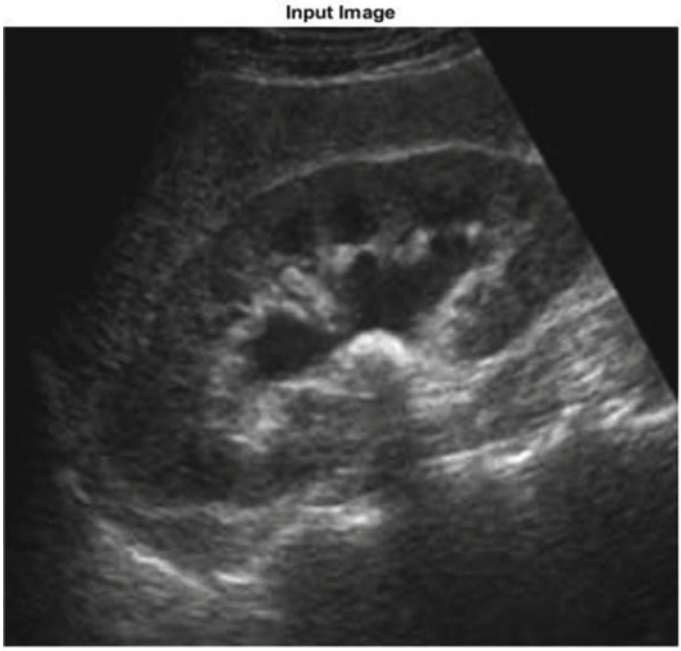


Fig. 57.2 Input image

properties are utilized for the segmentation of the active contours of the ultrasound image. Active contours have been generally utilized as appealing images segmentation techniques since they generally deliver sub-regions with constant margins or boundaries [7]. But in kernel-based edge detection methods like sobel edge detectors will generate discontinuous margins or boundaries. Application of level set hypothesis has given greater adaptability and comfort while execution of active contours. Morphological operations depend on the comparative demand of pixel values but not on the statistical values and consequently fit for the processing of all binary images. Morphological operations are then connected to grayscale images to unrevealing the light transfer functions and to get accurate pixel values. Then the final segmentation is done. Finally, extraction of the required GLCM features from the de-noised ultrasound image is finished. The original ultrasound image prior and then afterward preprocessing is as shown in Figs. 57.2 and 57.3.

57.2.4.2 Extraction Process of Ultrasound Image Features

One of the characterizing characteristics of texture is the spatial distribution of gray values. Based on statistical, structural, model-based and transform information, the texture features can be extracted by an excellent method by applying gray-level

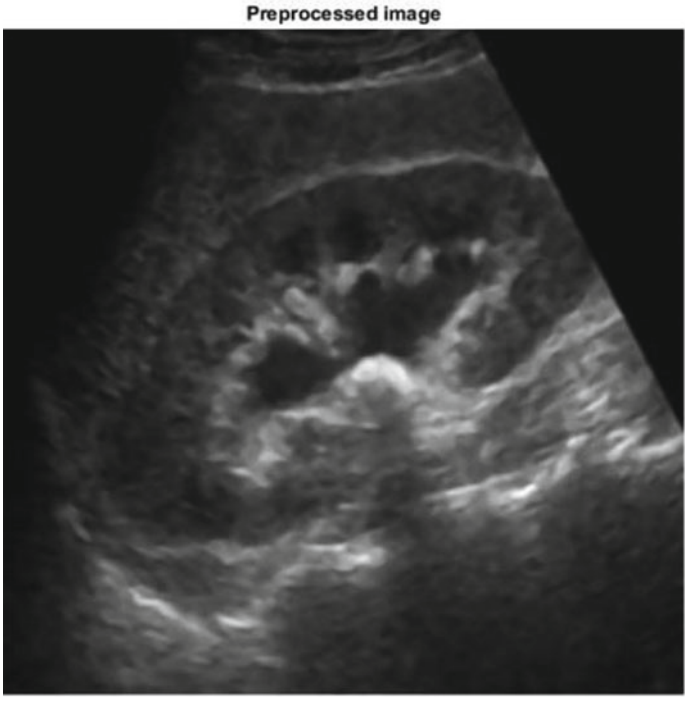


Fig. 57.3 Preprocessed image

co-occurrence matrix (GLCM) algorithm. GLCM contains statistical data of spatial relationship of the pixels in the image.

Table 57.1 portrays the 22 input parameters of the ultrasound which were utilized for training the neural network. Some such features are homogeneity, entropy, divergence, cluster color etc., as mentioned in Table 57.1. Homogeneity is the feature which measures the appropriate distance of the components in the GLCM to the GLCM diagonal. For diagonal GLCM, homogeneity is 1. Image entropy is an amount which is utilized to portray the measure of data which must be programmed by a compression algorithm. The perfectly flat images will have zero entropy. Divergence value is a measure that evaluates the reliance between the sequences. Cluster color feature is characterized by the skewness of the network (or) absence of the symmetry. The image is not symmetric, when the cluster color is high. Co-occurrence matrix will be peak around the mean values when cluster prominence is low. In the image, it implies that there is small distinction in the grayscales.

Table 57.1 GLCM features and their allowed values of ultrasound

S. No.	Features	Allowed values
1	Auto-correlation	650.8546
2	Contrast	300.2537
3	Correlation	0.0668
4	Correlation: [1 2]	0.0668
5	Cluster prominence	2.9658e+05
6	Cluster shade	164.2457
7	Dissimilarity	13.9130
8	Energy	4.7758e−04
9	Entropy	7.7256
10	Homogeneity	0.1397
11	Homogeneity: [2]	0.0689
12	Maximum probability	6.8378e−04
13	Sum of squares	811.4439
14	Sum average	50.3571
15	Sum entropy	2.5378e+03
16	Difference variance	4.4214
17	Difference entropy	378.7012
18	Information measure of correlation 1	3.6916
19	Information measure of correlation 2	−0.0030
20	Inverse difference is homom	0.1515
21	Inverse difference is normalized	0.7828
22	Inverse difference is moment normalized	0.8872

57.2.4.3 Analysis and Results

The developed GLCM feature extraction algorithm for extraction of features in ultrasound is implemented in MATLAB 8.5. Initially, the implemented algorithm peruses ultrasound image in a jpg format and performs preprocessing that it expels the noise from the input image by the way of padding. To concentrate ultrasound image features, the GLCM algorithm is utilized. Then the final segmentation is done by extracting the region properties [8]. The final segmentation is shown in Fig. 57.4.

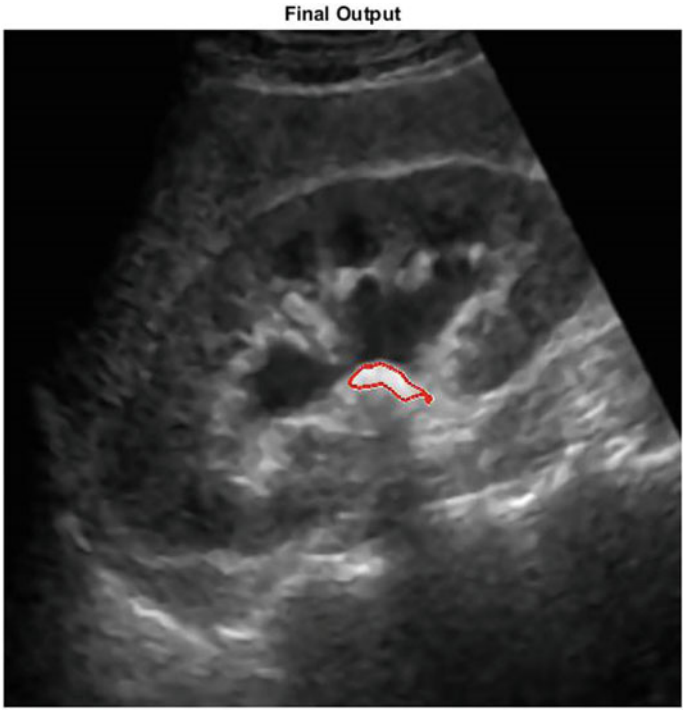


Fig. 57.4 Final segmented image

57.3 Predictions of Nephrolithiasis Based on Ultrasound Extracted Features Using Neural Networks

57.3.1 Introduction

In spite of the fact that the components are extricated consequently from ultrasound by the created GLCM calculation, still it is troublesome for a physician to diagnose nephrolithiasis. Since the extracted feature will not be same for all the patients, instead they vary from person to person based on age, sex, weight and height. This may bring inaccuracy of the diagnosis and may increase the time delay of the diagnosis [9]. Accordingly so as to provide a support system for the physician and to diminish the analysis time, it has turned into a requesting issue to build up an effective and solid clinical system [10, 11]. Since neural networks have indicated incredible potential to be connected in the advancement of clinical system for diagnosis of renal syndromes, in the present review additionally, a system is developed to perform the diagnosis of nephrolithiasis in view of the ultrasound features.

57.3.2 Parameters

The parameters that are utilized to perform diagnosis classification of nephrolithiasis in view of ultrasound are age, sex, weight, stature and the GLCM features extracted from the ultrasound image. In the present review, to prepare and test the neural system, an aggregate number of 50 samples are utilized, and then one output parameter is shown on which the experimental analysis is performed.

57.3.3 Procedure

In the present review, to build up a system for diagnosis of nephrolithiasis in view of ultrasound features, at first the ultrasound features, i.e., GLCM features are extracted by applying the GLCM feature extraction algorithm to the binary converted input images. Then, a feed-forward back-propagation neural network is built by taking 22 input parameters of the extracted features and the essential information.

57.4 Analysis and Results

The system is designed and executed utilizing MATLAB 8.5 by implementing the feed-forward back-propagation neural network model. In order to execute back-propagation network model, at first a feed-forward neural network is built with 22 (twenty-two) input nodes, 10 (ten) hidden nodes and 1 (one) output node. The input parameters utilized as a part of this system are the age, sex, weight, height, etc., and the GLCM features extracted from the ultrasound image and the output parameter demonstrate result of the diagnosis in terms of either normal or abnormal, i.e., if ultrasound of the kidney does not have any stone, then the condition is normal; if the kidney consists of a stone, then the result is abnormal. The feed-forward network is built with 22 (twenty-two) input nodes, 10 (ten) hidden nodes and 1 (one) output node. Construction of neural network with assigned nodes is shown in Fig. 57.5.

Training is done by utilizing the extracted GLCM features. The algorithms mentioned in the neural network train tool appeared underneath in Fig. 57.6. Here the data which are collected are separated randomly. Then that random data are used for training by using a learning method called Levenberg–Marquardt method. The obtained values between the gradient descent and the Gauss–Newton were adaptively differs by the Levenberg–Marquardt method.

$$\left[J^T W J + \lambda I\right] h_{lm} = J^T W\left(y - \hat{y}\right)$$

where minor estimations of algorithmic parameter 'λ' result Gauss–Newton update and vast estimations of 'λ' result in gradient descent update. The parameter 'λ' is

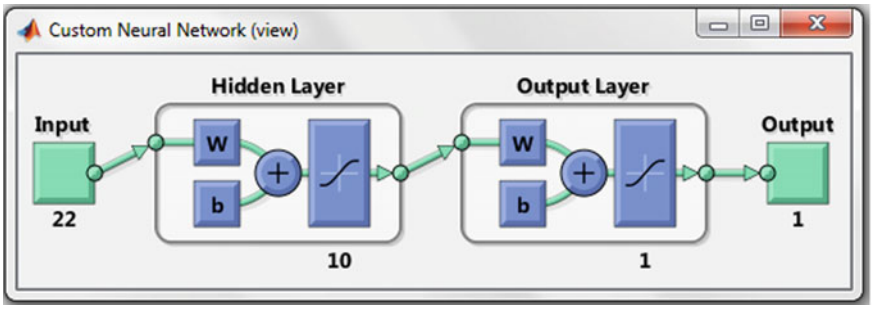


Fig. 57.5 Construction of feed-forward neural network with 22 input nodes, ten hidden nodes and one output node

modified to be extensively with the goal, which updates the small strides in the steepest-descent direction. If any iteration occurs which leads to worse calculation, then 'λ' is expanded. $\left(x^2(p + h_{lm}) > x^2(p)\right)$.

As the result progresses, λ is diminished. Levenberg–Marquardt strategy approaches the Gauss–Newton technique by which the result normally accelerates to the local minimum value.

Marquardt's refresh relationship

$$\left[J^T W J + \lambda\, diag\left(J^T W J\right)\right] h_{lm} = J^T w\left(y - \hat{y}\right)$$

The obtained values of 'λ' are normalized to the values of $J^T W_J$.

Then the performance is evaluated after the training is completed and then the values are generated. The performance is assessed by considering epochs, and mean square error is shown in Fig. 57.7. The graph appeared beneath portrays the performance of error at 6 epochs. The trained data are demonstrated with a blue line and the test data with a red line. The best validation is taken at the point where both train and test are equivalent. Here the best approval is at epoch 0, and the value of validation performance is 0.31456.

The plot appeared in Fig. 57.8 reveals the data of training and the validation of target and output classes. In the graph, the input is represented as data denoted by 'o'. At that point, the spotted line represents the linearity of result and the target. The fit esteem begins at 1.38 and increments marginally all through the training. The fit esteem does not change even when the target classes expanded. The validation graph shows the fit value linearly increments along with the target and output classes. Then for test data, the plots are represented below in Fig. 57.9.

Figure 57.10 reveals the data of gradient, mu and validation fail at 6 epochs. The gradient value follows a straight line, it continuously diminishes, and the esteem becomes plainly 0.00063259 at the sixth epoch. Then the mu value increments for a point of 1 and falls, and at the sixth epoch the value is 1e−08. The validation checks increment along with the expansion in the epochs, i.e., the validation checks and the epochs are linear and the validation check value at the sixth epoch is 6.

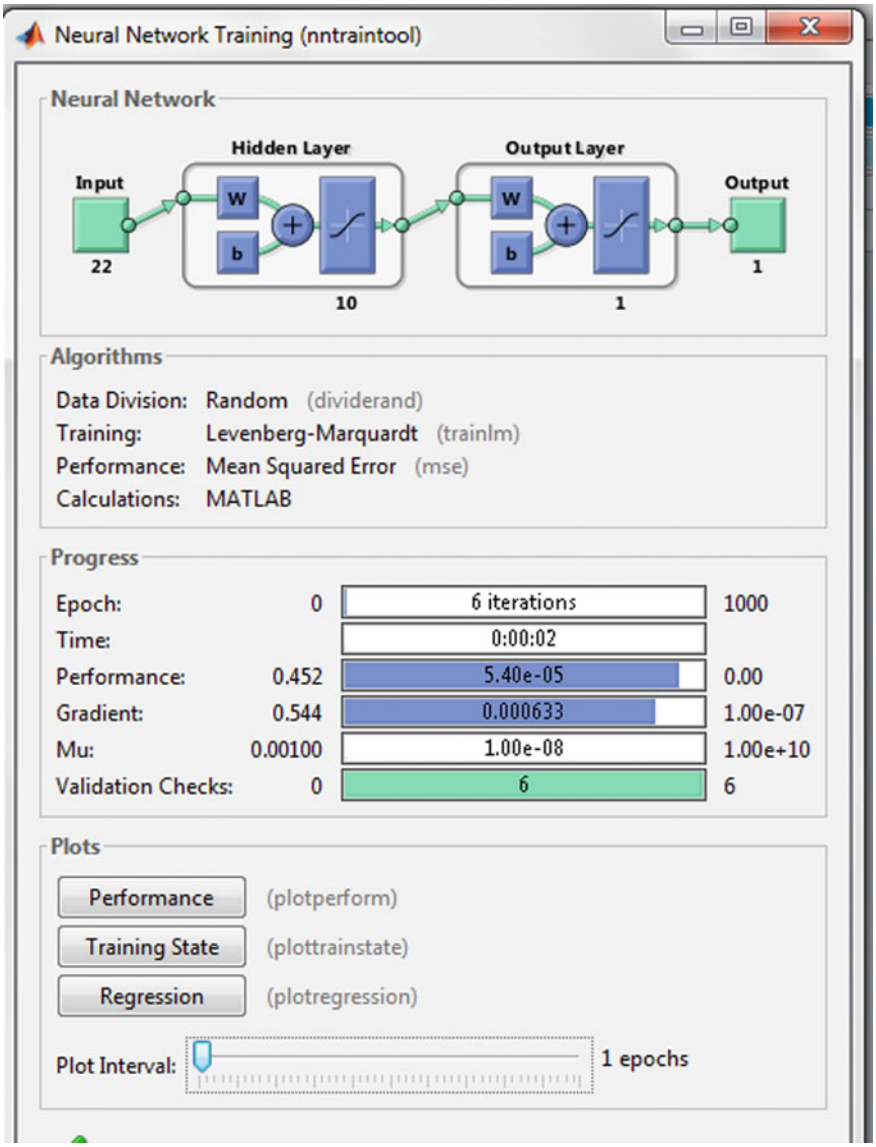


Fig. 57.6 Neural network training

57.5 Conclusion

The prediction of nephrolithiasis based on ultrasound components is a troublesome undertaking for nephrologists because of the reason that the ultrasound images will not be comparative for all individual rather it changes from individual to individual.

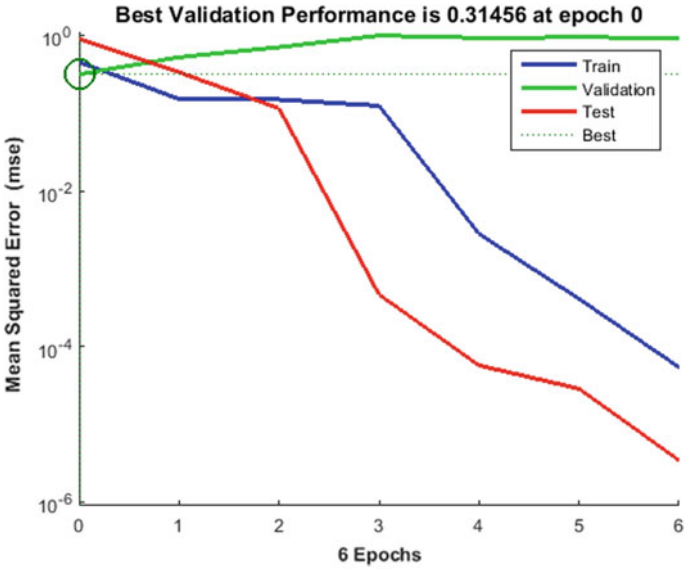


Fig. 57.7 Best validation performance at epoch 1

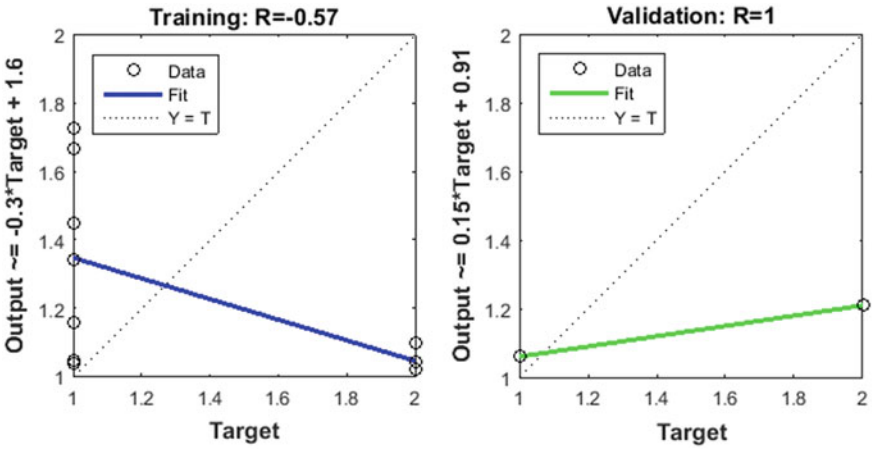


Fig. 57.8 Plot of training and validation for output with regard to target classes

Consequently, the systems can be utilized by a nephrologist to automatically extract ultrasound features and to perform the prediction of stone automatically based on the extracted features. The developed system diminishes the conclusion time and enhances the quality of the analysis. The extracted features are from an ordinary individual, and then the result, otherwise, gives the result with the final segmented image with a bouncing box precisely.

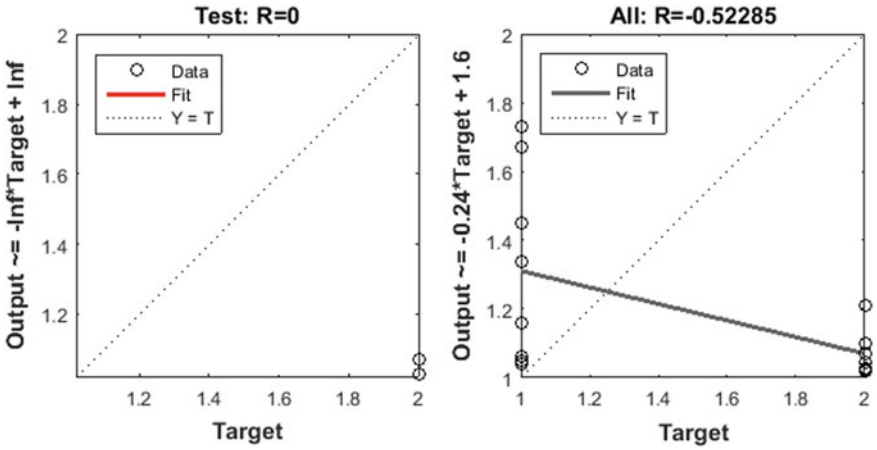


Fig. 57.9 Plot of testing data for output with regard to target classes

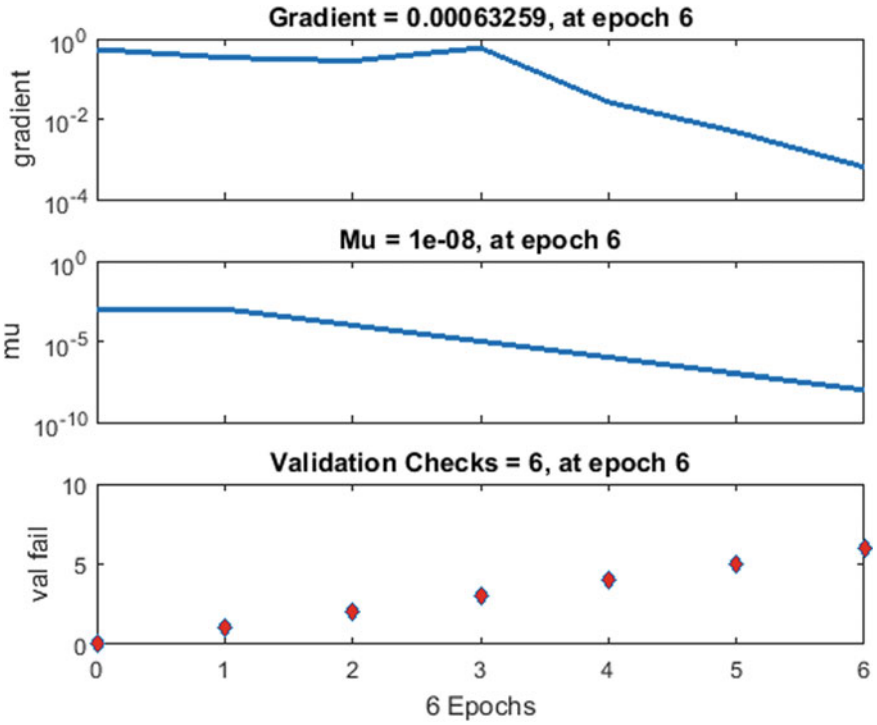


Fig. 57.10 Plot of training state at six epochs

References

1. Attia, M.W., El-Din Moustafa,H., Abou-hadi, F.E.Z., Mekky, N.: Classification of ultrasound kidney images using PCA and neural networks. Int. J. Adv. Computer Sci. Appl. (IJACSA) **6**(4) (2015)
2. Ramana, K.V., Korrapati, R.B.: Lung Tumor Segmentation via Fully Convolutional Neural Networks. Austin Ray, CS 231N, Winter 2016 (2016)
3. Neural network based classification and diagnosis of brain hemorrhages. Int. J. Artif. Intell. Expert Syst. (IJAE) **1**(2)
4. Francisco, M.P., Juan Manuel, G.C., Antonio, S.P., Daniel, R.F.: A robust model of the neuronal regulator of the lower urinary tract based on artificial neural networks. Neurocomputing **71**(4–6), 743–754 (2008)
5. Kalyan, K., Jakhia, B., Lele, R.D., Joshi, M., Chowdhary, A.: Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images. Adv. Bioinform. **2014**, 14, Article ID 708279 (2014)
6. Bayo, A.L.: Computer Aided renal calculi detection using convolutional neural networks. International Master's Thesis Studies from the Department of Technology at Örebro University
7. Viswanath, K., Gunasundari, R.: Analysis and Implementation of Kidney Stone Detection by Reaction Diffusion Level Set Segmentation Using Xilinx System Generator on FPGA, 10p. Research Article Hindawi Publishing Corporation VLSI Design Volume 2015, Article ID 581961
8. Dayhoff, J.E., DeLeo, J.M.: Artificial neural networks—opening the black box. Cancer **91**, 1615-1635 (2001)
9. Brause, R.W.: Medical analysis and diagnosis by neural networks. https://www.informatik.uni-frankfurt.de/asa/papers/ISMDA 2001k.pdf
10. Lisboa, P.J.G. : A review of evidence of health benefit from artificial neural networks in medical intervention. https://www.openclinical.org/docs/ext/lisboa2002.pdf
11. Zukić, D., Elsner, A., Avdagić, Z., Domik, G.: Neural networks in 3D medical scan visualization by June 2008

Chapter 58
Multi-resource Task Scheduling for Minimum Task Completion Time in Cloud Computing Using Credit-Based Assignment Problem



Animoni Nagaraju and Y. Rama Devi

Abstract Task scheduling process problem is a focus and challenging issue in the Cloud Computing. The Cloud enumerate resources skillful and gain the most profits with scheduling process is the ultimate objective in the one of the Cloud enumerate overhaul providers. In this paper, we used multi-resource credit-based assignment problem to assess the entire group of multi-resource task scheduling in task get in line and finding the minimal completion time of all tasks in the cloud enumerate. Here cost/time matrix has been created as the reasonable tendency of the task to be the assigned in multi resources. Here we concern Cloud enumerate is a latest new assessment method of delivering the technology in the consumer by means of Internet servers for dealing out and data storage, while the client process uses the data.

58.1 Introduction

The technology experts considered the cloud enumerate is hovering to modernize the technique we access tools which it's going to be the same as task accruing because the profit-oriented of the technology more than a decade ago. Cloud Enumerate empower innovation. It alleviates the necessity of innovators to seek out possessions to extend, test, and build their innovations presented to the customer community. Innovators are liberal to specialize in the modernization instead of the logistics of discovery and organization resources with the purpose of enabling its innovation. In order that the task development evils and source administration are transmit to

A. Nagaraju (✉)
Department of CSE, University College of Engineering, Osmania University, Hyderabad, India

MallaReddy Institute of Technology and Science, Hyderabd, Telangana, India

Y. Rama Devi
Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, India
e-mail: yrd@cbit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_58

competence of entire cloud enumerate facilities. These tasks are matching process on the junction of the cluster through strategy which makes every effort to stay the job as on the brink of the info as potential. The scheduling algorithms within scattered processes usually contain goals of allocation consignment on processors and its maximizing theirs utilizations while minimize the total assignment effecting time more than a few heuristic approaches have been introduced during assignment scheduling. The inspiration of this paper is the work out a scheduling method which follows the Lexi—search approach of A. Nagaraju and V. V. Haragopal to seek out an optimal best feasible jobs. Assignment scheduling has been treated as general obligation problem to hunt out the minimal charge. In this, the cost matrix are generated from the probabilistic factor supported a quantity of most significant situation of efficient assignment scheduling just like the foremost important task interval during a resource. The value for assigning task into a source is probabilistic outcome considering.

In this context, a number of contributions are established for optimizing their performance more. The resource scheduling and the workload balancing is the similar directional effort for optimizing the computational performance of the cloud infrastructures. The proposed work is meant to live the performance of cloud scheduling approaches that are claimed to optimize the presentation of the cloud computing.

Cloud Enumerate means both applications and services are moved into the web (cloud). Cloud enumerate provides the users with any quite software or hardware services through the web on the idea of pay as you employ. Cloud enumerate is define as comparable and scattered process which comprise of group of interconnected and virtualizes computers that be provisioned dynamically and existing together or more combined enumerate resources supported service levels agreement (SLA) established from beginning to end mediation between cloud services providers and users [1]. Cloud enumerate delivers transportation as a service (IaaS), Platform as a service (PaaS), and Software as a service (SaaS) [2]. IaaS provides the users with infrastructure in sort of a virtual Machines (VM) to run any applications. PaaS provides an application development platform to users or developers to develop their own cloud applications and SaaS allows users to run existing software applications. Cloud enumerate is rapidly developing Computation model which has moved the management of hardware, software and other enumerate resources from users to cloud service providers.

There are numerous issues regarding cloud computing data security, energy conservation, service availability, expandable storage management, task scheduling. But task scheduling is usually the foremost topic of research in cloud enumerate [3]. the first objectives of this paper is to reinforce the scheduling policies. In cloud enumerate many tasks needed to be executed from the resources, available so on realize high performances, optimal completions time, reduce response times, and effective resources utilizations [4]. Due to these different objectives, there's a requirement to develop and propose a scheduling algorithm which will be employed by task scheduler to appropriately allocate tasks to resources.

Cloud Enumerate provides the prospect to employ enumerate resources more the web exclusive of owning the infrastructures. The most contents of Cloud Enumerate

is to the handle software appliance, data storages, and processing facility its are assign to peripheral user on the insist all the way through the web and give just in support of come again its they are use. Tasks arrangement in cloud enumerate is that biggest challenge, since it many tasks got to be executed by available resource so as to satisfy user's requirements. To realize the most excellent performance, minimize the total completion time, reduce reaction times, and maximizes assets utilizations there's got to concentrate on these challenges.

From this paper, we used multi-resource credits-based assignment problem to evaluates the entire group of multi-resource assignment scheduling in task wait in line and finds minimal completions time of all tasks in the cloud computing. Now cost/time matrix have been generated as fair tendencies of a task to be assigned in multi possessions. Here we are appropriate Cloud enumerate is the latest new enumerate method of distributing technology to the consumers by use Internet servers for dealing out and data storages, while this client process uses the data. *The main idea of the planned algorithm is to assign responsibilities to assets properly so as to realize an resourceful weight matching and reduce completions time. The investigational outcome indicates the proposed approach by lexicographic search.*

58.1.1 Task Scheduling in Cloud Computing

Arrangement is mentioned to be as a set of policies for organization the ordered of work to be executed by a enumerate process [7]. The task arrangement in Cloud data center are a set of commands and elements that control and decide on the assignment to be executed on the available assets between a group of feasible everyday jobs at a specific time [8]. In Cloud data centers, the task arrangement algorithms are responsible for assign the farm duties submit by the user to accessible assets. The main benefit of task scheduling algorithms is to realize a highly performance enumerate and therefore the most excellent process throughput [2, 7, 9]. Schedulings lead the CPU memory and to realize maximum source utilization, require good scheduling policies.

In Cloud Enumerate location, responsibilities are submitted to info Center Brokers by the end users. This Data Centered agent is an agent between a Cloud Users and providers and the liable for arrangement farm duties on Virtuals machines (VM). Data hub may be virtual communications for shelter possessions and consisting of variety of Hosts. These submitted tasks are programmed consistent with the arrangement policy employed by the info hub Brokers. Broker communicates frankly with a Cloud manager and assigning tasks to Virtual machines within the Host info Center [10].

There are several types of scheduling as maintained by to unlike policy such that 'preemptive and non- preemptive' arrangement, statics and dynamic arrangement, instant and group setting up, integrate, and spread development [4]. Scheduling of tasks are taken into account because the process of choosing the simplest source accessible for tasks execution. Task scheduling approaches aim at minimize the total completion time of task and maximize source consumption so as to satisfy customer

requirements, below diagram represents the proposed architecture, present paper concentrates on scheduler manager.

58.2 Proposed Methodology

An algorithm is extended based on lexicographic search algorithm. From this approach, every answer to the problem is defined as a word and solutions are to be produced in a quantity of hierarchy of their values. Each a partial word defines 'a block of words with itself as leader and is bound on the value of objective function for this block of words is considered'. If this 'bound is more than a value of objective function for the known feasible assignments', then the entire block of words is discarded. Algorithm start with 'feasible assignment' this provides a reasonably excellent upper bounds. Articulate on the values of the objective function. The initial possible assignments is obtained using a heuristics come within reach of in scheduling theory designed for dealing out tasks on matching machines minimizing the makespan approach 'examines the optimality' of this most feasible assignment and if found non-optimal, proceeds to find an optimal feasible assignment in a restricted number of iterations The above methodology explains "Lexicography" means "dictionary compiling". By "lexicographic arrangements" we mean the arrangement of the given letters or numbers as in the dictionary, Searching for a particular arrangement in this ordered, set is similar to searching for a word in the dictionary. The lexicographic search approach was developed by pandit SNN (1962) in the context of loading problem and was subsequently employed for many other combinatorial programming problems. The method adopted by pandit SNN is as follows: The set of items is partitioned into two subsets P and Q where P is the selected set and Q is the rejected set. For every word in P, i.e., for every subset of the selected set, a feasible and acceptable exchange is obtained, then the sets P and Q will be adjusted accordingly and the search will be started again. If such an exchange is not available, then the current selection is the optimum solution. Since the items in the sets P and Q are arranged in a standard order, a monotonic relationship exists between the values of the words belonging to partially ordered subsets from P and Q, which is helpful as a stopping bound criterion for the search in the set B. The present discussion is confined to the application of lexicographic search in the context of Multi-Resource Credit-Based Assignment Problem with minimizing objective function.

58.3 Algorithm of Multi-Resource Credit-Based Assignment Problem with *n* Tasks

Step 1: We consider the Weighted multiple resources it's rehabilitated in to two appliance n task assignment problem by using weighted average technique. i.e. various

clusters to collect the data the cloud enumerate in the prairies based on the probabilistically approaches apply to use weighted multi resources it constructed to credit-based assignment problem.

Step 2: Two appliance with n tasks problem, we consider likely number of permissible acceptable selections used for appliance(machine) 1 and a number of acceptable selections for appliance(machine) 2, out of the which 2 appliance(machines), there will be' n task selection's, and its selection are- to be computed.

Step 3: Now we constructed alphabets plunk, which be an arrangements of times in a growing array of the problem for appliances 1st and 2nd machines. Here chosen the arranging of tasks is to done with indexed number by the not breaking of original order sequence in the tasks.

Step 4: From considering the changeable combinations of 'the at least' and 'the at most' constraint we get sub-problems for given 2 appliance assignment problems, now initially get the trials solution meant for the 2 machines problem.

Step 5: Now, lexi search methodology performs, scientifically we make 'the incompletes of words as of the search table with the new-labels, used for this every machines, except proceedings innovative task names considering that well be accrue time integrated in a word so far, and also the bounds for remaining part of partial word i.e. bounds for every realistic word's in lexical block, for its the current incompletes block is the leader'. If occurred bound is superior than the trial solution value on hand, this leader are discarded and the next incompletes word of the same length or it's next super block are head, as the holder may be the selected on the present incomplete word. Above steps record in search bench accessible in below.

Step 6: Using a step 4, we fixing the objective in algorithm Step 5, to achieve the optimum solution for expected objectives. i.e. (Minimize total accumulating machines time, this true for 2, 3 and m machines).

Step 7: From the above steps, we get the best feasible solution is evaluated for comparing with trial solutions from optimal solutions of step (6) in the various subproblems of step (4), from this we obtain best optimal solution in this step.

Now, we are implementing above algorithm in the 2 appliance and 7 Tasks problem is define problem in the next section.

58.4 Lexi Search Procedure of Multiple Resources Credit-Based Two Appliance *n* Task Assignment Problem

Numerical Illustration Discussed by Using Lexi Search Procedure for MR Credit-Based AP

In one byword, a methodology, processing, and apparatuses provide prioritized, faired, and weighted assignment scheduling for collective execution stand in line. New tasks are prioritized accordingly to their process in the requirements and held in one or more queue according to their precedence levels. Tasks are retrieved from the

each queue in fair and weighted approach according to the priority level of each wait in line. In one of the embodiment, tasks of approximately equivalent total processing time are rived from each storages queue in the given task retrievals cycle.

In the next detailed description of the various embodiments of invention, numerous specifics detail are set forth in order to be provided thorough understanding of the various aspects of one or more embodiment of the inventions. However, one or more embodiments of the inventions may be practiced without these are specific details. In other the instances, well-known methodologies, procedure, and/or component has not been described in detail so as not to unnecessarily obscures aspects of embodiment of the invention. In the next description, certain terminology is used to describes certain features of the one or more embodiments of invention. For instance, the 'frame include any block or arrangements of data or informations. The terms "informations' is defined as voice, data, address, and/or control'. The term "task identifier" includes any depiction corresponding to a unique numbers frame. One part of an embodiment of inventions provides a methodology, Process, and apparatus having multiple executions in the queues of different priority coupled to a shared in the execution queue to provide the shared executions queue with a available processing tasks. Another aspects of an embodiments of the inventions provides fair, exact and weighted execution of the prioritized processing tasks.

Combinations of data Coating by the MCE with weighted linear combinations, each of the coating is assigned a weights, or the level of significances, that is determines to its influence on the layer producing as results of the combinations. In this current study weighting were assign by based on knowledge of element concerned as well as the consensus obtained as of the focus groups weightings is assigned so, that while they are combined to have total value. The combination of three layers considered the probabilities layer $1 = 0.5$, layer $2 = 0.25$, layer $3 = 0.25$, This value assigned to the each cell in the each data layers is then multiplied by the weights. Corresponds cell's in each the layer is then added to together produces a final combined assignment table. The example of the processes using the 3 layers and the weightings describes above is given below figure considered multi-resource credit-based two appliance n task assignment problem.

We have generated the problem by using above procedure from various clusters to collect the data to use multi-resource constructed to credit-based assignment problem to calculate the entire group of multi-resource task scheduling in assignment queue, now we find minimal completion times of all tasks assignment from the cloud enumerate in the prairies based on the probabilistically approaches. From this cost/time matrix has been generated, as the fair tendency of a task to be assigned in multi resources. Now we apply Cloud enumerate is the latest new enumerated method of dispatch technology to the consumer by via Internet servers for processing and data storage, while the client process uses the data.

After construction of the credit-based assignment problem, we used the lexi search approach to make the task assignment scheduling in the processatical order.

Considered from Table 58.1: Observed this existed the credit-based assignment problem for 2 appliance 7 tasks: In next fragment, we are demonstrating algorithm

through an illustration: regard as the following example 2 appliance 7 tasks situation is shown in Table 58.2.

For this problem, according the algorithm make alphabet-table, which are an array of "times" in growing order of the problem from 1st appliance and second appliance. The arrangements of tasks are complete with an index numbers by the not breaking original order sequence of these tasks and shown in Table 58.3.

Table 58.1 Example of the problem multi-resource credit-based two appliance 7 task assignment problems

		M1	M2														
	J1	5	3		J1	2.5	1.5										
	J2	1	3		J2	0.5	1.5										
Layer1	J3	2	4	x 0.5	J3	1	2										
	J4	4	5		J4	2	2.5										
	J5	2	2		J5	1	1										
	J6	7	6		J6	3.5	3										
	J7	3	4		J7	1.5	2										
		M1	M2			M1	M2										
	J1	4	7		J1	1	1.75		J1	4.25	4.5		J1	4	5		
	J2	1	8		J2	0.25	2		J2	2.25	4	Layer are then rounded to give whole numbers	J2	2	4		
Layer2	J3	5	4	x 0.25	J3	1.25	1	Layer added together	J3	3.75	4.25		J3	4	4		
	J4	3	5		J4	0.75	1.25		J4	3.5	2.5		J4	4	3		
	J5	4	7		J5	1	1.75		J5	3.5	4.75		J5	4	5		
	J6	7	6		J6	1.75	1.5		J6	6.25	5.75		J6	6	6		
	J7	3	5		J7	0.75	1.25		J7	3.25	5		J7	3	5		
		M1	M2			M1	M2										
	J1	3	5		J1	0.75	1.25										
	J2	6	2		J2	1.5	0.5										
Layer2	J3	6	5		J3	1.5	1.25										
	J4	3	4	x 0.25	J4	0.75	1										
	J5	6	2		J5	1.5	0.5										
	J6	4	5		J6	1	1.25										
	J7	4	7		J7	1	1.75										

Table 58.2 2 appliance 7 tasks

	1	2	3	4	5	6	7
M1	4	2	4	4	4	6	3
M2	5	4	4	3	5	6	5

Table 58.3 Alphabet table

S. NO.	M1: *ti*	M1: *ji*	M2: *ti*	M2: *ji*
1	2	2	3	4
2	3	7	4	2
3	4	1	4	3
4	4	3	5	1
5	4	4	5	5
6	4	5	5	7
7	6	6	6	6

Now, we fix the least and at most task (jobs) constraint on this problem with 7 task selections, we considered the fixed and is taken on first appliance as 2 & 4 tasks are to be processed. Since the requirement is that exactly on second appliance 3 & 5 tasks are to be processed on the apparatus M1 & M2 correspondingly, a complete lower bound meant for a few of the feasible assignment get as sum of the only least 2 and any least 5 processing's time in 2 machines, i.e., total of 1st two and 1st five timing's on these machines (appliance) is considered, that is for the problem the sum**:{5,6,}{4,2,3,1,7,},10 + 21 = 31 is bound by the Arora optimality.**

While timings on two apparatus are the independent of separate tasks allotment, bound settings can be completed, for each appliance separately in the parallel or the one can be first computing assignments on appliance 1 (say) and then we go to appliance2, enumerate bound section for an unassigned part on the 1st appliance explicitly, keeping the simpler (this is the relatively very less efficient bound, is the calculated once for the all, for M2, irrespective of tasks already assign for the M1). Also, the computationally, simpler bound is not accumulate, the number of tasks, yet to be by now allotted to on the machines, but just to reproduce the number by the 'next' processing time in the alphabet table.

In this following example, allocation is lexicographically prepared first on appliance Machine1, with constant minimum of bound for Machine 2, and Machine 1 allotment which is finished, and then only to go for more the exact bounds for components based on the Machine 2. "The search desk analytically 'Generates' partial words via the latest labels, for the each of machine, but the proceedings are original task numbers as well as the accumulates times integrated in the word, similarly so far, and its also the bounds meant for remain part of partial word, this bounds, for all the feasible words is in the lexical blocks, for which this current, incomplete block is a leader. If its bound are greater than trial solution significance on hand, its leader is redundant, the subsequently incomplete word of the same length, or it's next super block is a leader, as the similar case may be selected on current incomplete word". Above steps are recorded in the search table existing below; this construction of search slab, as explained over is illustrated below: Search table gives for minimization Objective by lexi search approach.

In the search table using the lexi search approach by C program we obtain the task constraints allotted to the two machines (i.e. M1, M2) for the constraints considered at least and atmost tasks are first appliance selections J1:J2, J3:J7. And second appliance selections J1:J3, J4:J7. Occurred by considered constraints. From the constraint J1:J2, J3:J7, we obtain the optimal solution by lexi search approach is {2,7,}{4,3,1,5,6,},cost/time is $5 + 23 = 28$ and from the constraint J1:J3, J4:J7,**we obtain the optimal solution by lexi search approach is: {2,7,5,}{4,3,1,6,}, cost/time is 9 + 18 = 27**.

From the above two constraints, we obtain best optimal solution is MIN(MIN), we obtained the Selections (M1):3,(M2):4 are 2,7,5 on appliance one and 4,3,1,6 are on appliance two. **The best optimum solution optimal cost/time Is 27**.

From the above algorithm, we obtain the comparison table for above problem between constrains with objectives in Table 58.4.

Table 58.4 Optimal solution comparison table

	Constraints	Lexi search algorithm	Aurora algorithm (trial solution)
Objective	(M1):2 (M2):5	$5 + 23 = 28$	$10 + 21 = 31$
	(M1):3 (M2):4	$9 + 18 = 27$	
Best optimal solution	(M1):3 (M2):4	$9 + 18 = 27$	

58.5 Computational Experience

This idea is implemented on randomly generated 100 problems of $m = 2$ machines and $n = 20$ tasks for which M1 = the total appliance time on machine 1 and M2 = the total time on machine 2 are computed and also minimum machines accumulated times computed. With the above-cited constraints is verified and the results are tabulated, here 10 problems result are tabulated in Table 58.2. The data shows the pictorially in the Fig. 58.1 revels that the minimum time and it shows bound line on machine 1 and machine 2 This we observe minimum total time for accumulated appliance times is very less. The optimum solutions of Multi-resource task scheduling for minimum task completion time in cloud enumerate using the credit-based assignment problems size 20 obtained. Which is best probabilistically expectations in the data selections with best task Scheduling in the cloud computing table given in Table 58.5.

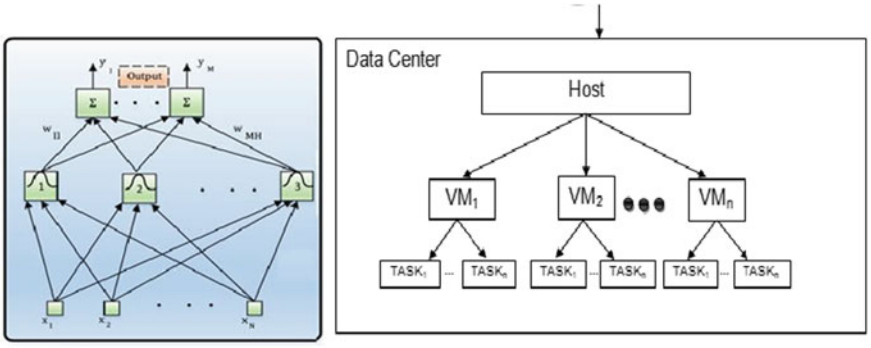


Fig. 58.1 Multi-resource task scheduling structure

Table 58.5 Optimal solutions for different problems

S. NO.	M1 (total time)	M2 (total time)	Total time (M1 + M2)
1	119	218	337
2	146	186	332
3	110	141	251
4	110	119	229
5	121	140	261
6	152	171	323
7	135	165	300
8	187	187	374
9	154	254	408
10	136	176	312

58.6 Conclusion

With encroachment of Cloud technology fast development, present are new needs for apparatus to learning, analyze the profit of the technology to implement and how the best to apply technology to large-scaled application The proposed method considers the arrangement problem as Multi-resource task scheduling for minimum task completion time in cloud enumerate using credit-based task problem.

In the cloud data center the data selections, the judgment assigning to the each cell in the each data layer are then multiplied by weights with probability. consequent cells in the each layer are to be then added together to produces a final combined task problem, in this, we apply lexi search approach to obtained best optimal task scheduling for multi resources for minimum task completion time in the entire cloud, that we compared with 100 randomly generated problems. That is most helpful to the data selections in probabilistically with expected results. This is new design of the problem which can focus in the extension of assignment problem to reach multiple clouds data selections with minimum time. Numerical demonstration indicate a good measures of success in the performances of the algorithm. The computational observation is that size of cloud increases the time for optimality increases as verified on computer 512 MB RAM.

References

1. Singh, Raja Manish, Paul, Sanchita and Kumar, Abhishek. *Task Scheduling in Cloud computing: Review.* 6, International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5, pp.No. 40–44 0975- 9646 (2014).
2. Lakra, Atul Vikas and Yadav., Dharmendra Kumar. *Multi-Objective Tasks Scheduling Algorithm for Cloud Computing Throughput Optimization*, International Conference on Intelligent Computing, Communication & Convergence, pp. 107–113 (2015).

3. Saxena, D., Chauhan, R.K., Ramesh, K.: Dynamic fair priority optiumization task scheduling algorithm in cloud computing: concepts and implementations. I. J Computer Network and Information Security, pp. 41–48 (2016).
4. Nagaraju, A., Hara Gopal, V.V., Pandit, S.N.N.: Time minimization assignment problem with variant objectives—a lexi search approach. Int. J. Agric. Statistical Sci. 6(I), 87–98 (2010)
5. Mohammadi, F., Jamali, S., Bekravi, M.: Survey on job scheduling algorithms in cloud computing. Int. J. Emerging Trends Technol.Computer Sci. 151–154 (2014)
6. Thomas, A., Krishnalal, G., Raj Jagathy, V.P.: Credit based scheduling algorithm in cloud computing environment. In: B.V. International Conference on Information and Communication Technologies, pp. 913–920 (2015)
7. Agarwal, A.. Jain, S.: Efficient optimal algorithm of task scheduling in cloud computing environment. Int. J. Computer Trends Technol. (IJCTT) **9**, 344–349 (2014)
8. Er-Raji, N., Benabbou, F., Eddaoui, A.: Task scheduling algorithms in the cloud computing environment: survey and solutions. Int. J. Adav. Res. Computer Sci. Software Eng. **6**, 604–608 (2016)
9. Bhavisha, K., Bhumi, M.: Review on max-min task scheduling algorithm for cloud computing. J. Emerg. Technol. Innov. Res. (JETIR) **2**, 781–784 (2015)
10. Himani, H., Harmanbir, S.S.: Comparative analysis of scheduling algorithms of cloudsim in cloud computing. Int. J. Computer Appl. **97** (2014)
11. Dehkordi, S.T., Bardsiri, V.K.: TASA: a new task scheduling algorithm in cloud computing. J. Adv. Computer Eng. Technol. **1**, 26–32 (2015)
12. Reda, N.M.: An improved sufferage meta-task scheduling algorithm in grid computing systems. Int. J. Adv. Res **3** (2015)
13. Kokilavani, T., Amalarethinam, D.I.G.: Load balanced min-min algorithm for static meta-task scheduling in grid computing. Int. J. Computer Appl. **20** (2011)
14. Patel, G., Mehta, R., Bhoi, U.: Enhanced load balanced min-min algorithm for static meta task scheduling in cloud computing. In: 3rd International Conference on Recent Trends in Computing, pp. 545- 553. Elsevier, Amsterdam (2015)
15. Rajeshkannan, R., Aramudhan, M.: Comparative study of load balancing algorithms in cloud computing environment. Indian J. Sci. Technol. **9**, 1–6 (2016)
16. Gokilavani, M., Selvi, S., Udhayakumar, C.: A survey on resource allocation and task scheduling algorithms in cloud computing environment. Int. J. Eng. Innov. Technol. (IJEIT) 173–178(2013)
17. Nagaraju, A., Rama Devi, Y.: Task Scheduling in Cloud Computing using credit based cluster travelling salesman problem. IOSR J. Eng. (IOSRJEN) **08**(10), 25–31 (2018) (2018)
18. Nagaraju, A., Rama Devi, Y.: Task scheduling in cloud computing using credit based variant of time dependent time minimization assignment problem. Int J Anal Experiment Modal Anal. **XII**(I), pp 307–313 (2020)
19. Panda, S.K., Pande, S.K., Das. S.: Task Partitioning scheduling algorithms for heterogeneous multi-cloud environment. Arab. J. Sci. Eng. **43**(2), 913–933 (2018)

Chapter 59
Insights into the Advancements of Artificial Intelligence and Machine Learning, the Present State of Art, and Future Prospects: Seven Decades of Digital Revolution

Prabhleen Bindra, Meghana Kshirsagar, Conor Ryan, Gauri Vaidya, Krishn Kumar Gupt, and Vivek Kshirsagar

Abstract The desire of human intelligence to surpass its potential has triggered the emergence of artificial intelligence and machine learning. Over the last seven decades, these terms have gained much prominence in the digital arena due to its wide adoption of techniques for designing affluent industry-enabled solutions. In this comprehensive survey on artificial intelligence, the authors provide insights from the evolution of machine learning and artificial intelligence to the present state of art and how the technology in future can be exploited to yield solutions to some of the challenging global problems. The discussion centers around successful deployment of diverse use cases for the present state of affairs. The rising interest among researchers and practitioners led to the unfolding of AI into many popular subfields as we know today. Through the course of this research article, the authors provide brief highlights about techniques for supervised as well as unsupervised learning. AI has paved the way to accomplish cutting-edge research in complex competitive domains ranging from autonomous driving, climate change, cyber-physical security systems, to healthcare diagnostics. The study concludes by depicting the growing share in market revenues from artificial intelligence-powered products and the forecasted billions of dollars worth of market shares ahead in the coming decade.

P. Bindra · G. Vaidya · V. Kshirsagar
Government College of Engineering, Aurangabad, Maharashtra, India
e-mail: vivek@geca.ac.in

M. Kshirsagar (✉) · C. Ryan
Biocomputing and Developmental Systems Lab, Lero, the Science Foundation Ireland Research Centre for Software, University of Limerick, Limerick, Ireland
e-mail: Meghana.Kshirsagar@ul.ie

C. Ryan
e-mail: Conor.Ryan@ul.ie

K. K. Gupt
Limerick Institute of Technology, Limerick, Ireland

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_59

59.1 Introduction

With an ultimate aim to surpass human capabilities, the technology is continuously evolving through the difficulties and has achieved greater milestones on their way [1]. Huge amounts of digital data and the need to keep abreast with the ever-evolving optimized versions of hardwares have fueled the rising interest of AI and ML [2]. Mainstream ML techniques such as classification, regression, dimensionality reduction, object recognition, and language processing are widely used in predictions, recommendations, detection, and many other similar applications. These technologies have already gained strong roots since the last decade and are expecting an exponential rise in the coming future [3].

The foreground of this in-depth analysis revolves around the past, present, and future of this entire paradigm with their current opportunities and future prospects. We continue with the present technologies under each type of ML and AI, their industrial applications and market trends followed by exhibition of the predictions and capabilities of the revenue market around their applications in the next decade, concluding the comprehensive study.

59.2 The Evolving Hype of AI/ML

This section highlights the history of evolution of AI, ML technologies, and their recent advancements by citing landmarks that led to the successful deployment of the wide applications this technology offers.

59.2.1 The Past

The term AI was first discussed in a context when a study was carried out jointly by three universities Dartmouth, Harvard, and Bell telephone laboratories which subsequently led to the birth of AI and soon followed the first programming language Lisp [4] for the AI researcher community. The machine learning paradigm was one of its kinds which learnt from experiences like a human and could solve problems by manipulating sentences in formal languages. The early 1960s witnessed the first-ever industrial robot 'Unimate' working in general motor's assembly lines. The mid-sixties gave birth to the chatbot 'ELIZA' [5] that facilitated dialogs between machines and humans (Fig. 59.1).

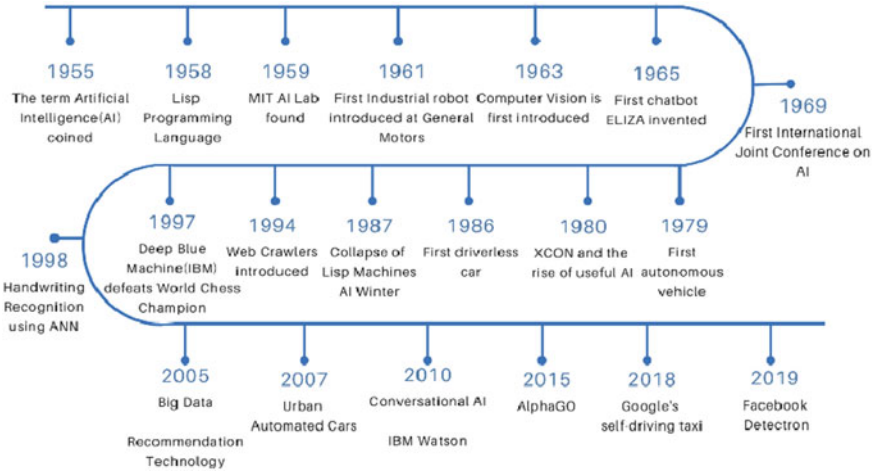


Fig. 59.1 Evolution of AI [6]

59.2.2 The Present

The late seventies led to the first autonomous vehicle 'Stanford cart' [7]. With the commencement of the 1980s, AI capacity at automation was explored with the Xcon program [8]. The mid-eighties saw the first Mercedes-Benz driverless car running on the empty streets of Munich. In the twenty-first century, the introduction of recommender systems designed to assist users in product selection choices based on their requirements led to the acceleration of the e-commerce market. IBM Watson [9], a real-time question–answer-based system was a technological revolutionary breakthrough in the last decade.

59.3 Subdomains of AI/ML

This section discusses the various emerging subdomains of AI and ML that are pioneers to the upcoming techniques in the field.

59.3.1 Artificial Neural Networks (ANN)

ANN [10] trains machines to solve problems in a way as a human brain does. The first neural network perceptron was developed by Frank Rosenblatt in the late nineteen-fifties. ANNs like convolutional neural networks, long short-term memory, recurrent neural networks, and auto-encoders gradually evolved over the years. These

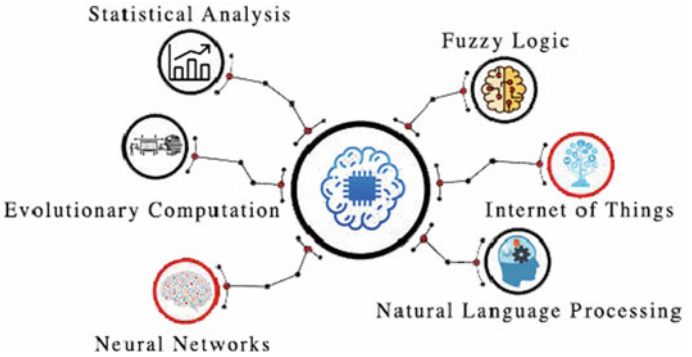


Fig. 59.2 Emerging subfields of AI and ML

techniques culminated into the first successful ANN-based handwriting recognition system in the mid-nineties. Applications of ANN include biometrics like iris, fingerprints, and face recognition systems which are now deployed across organizations that primarily require unique user authentication (Fig. 59.2).

59.3.2 Evolutionary Computation

A technique inspired by Darwinian evolutionary concepts of reproduction, mutation, recombination, and selection, introduced a new sub-domain in AI, known as evolutionary computation [11]. The underlying techniques like genetic programming and grammatical evolution [12] optimize solutions for enhancing computational efficiency.

59.3.3 Natural Language Processing (NLP)

Human–computer interaction through direct communication with devices like voice assistants and chatbots marked the NLP era [13]. Recent applications of NLP include answer sheet evaluation for examinations, text search and filtering, and lexical analysis.

59.3.4 Bayesian Networks and Fuzzy Logic

Statistical tools to incorporate cause-and-effect relationships in modeling complex systems came into being through the Bayesian networks [14]. They are impactful

within the broader armamentarium of AI methods with a wide array of applications in pharmacogenomics and physiologically based pharmacokinetic modeling.

The challenge of modeling human reasoning in an environment without measurements, vagueness, and uncertainty brought about fuzzy logic into existence [15]. The logic can be embedded from microcontrollers to large networked or workstation-based systems.

59.3.5 Internet of Things (IoT)

The advent of IoT [16] following the World Wide Web (WWW) created the digitized interconnected virtual and the physical world we live in today. IoT frameworks spanning organizations and countries have made it possible to integrate data from diverse sources leading to the development of smart cities across the globe.

59.4 Popular AI Methods

This section provides a brief insight discussing all the algorithms along with their suitability with respect to creating AI models.

59.4.1 Supervised Learning

Linear, multiple, and logistic regression analyses are used for modeling the relationship between predictor and output variables. Ordinary least square regression (OLSR) is used for estimating unknown parameters. Multivariate adaptive regression splines (MARS) gives near optimal solutions by allowing automatic selection of variables, interaction between predictors, handling missing values, and avoids overfitting with self-tests. Locally estimated scatter plot smoothing (LOESS) and locally weighted scatter plot smoothing (LOWESS) regression are used to reveal trends and cycles in input data by automatically choosing a smoothing parameter for a smooth curve. Ridge regression deals with multicollinearity in multiple regression [17] (Fig. 59.3).

Dimensionality reduction transforms data from high-dimensional space to low-dimensional space, hence retaining meaningful information. Techniques supporting this include principal component analysis (PCA) for unsupervised learning, linear discriminant analysis (LDA) in linear space, and quadratic discriminant analysis (QDA) for quadratic space for supervised learning, whereas partial least squares regression (PLSR) and principal component regression (PCR) for regression analysis [18].

Ensemble methods combine various multiple learning algorithms to stabilize the variance of predicted solution and thus increase the accuracy. Bagging is based on the

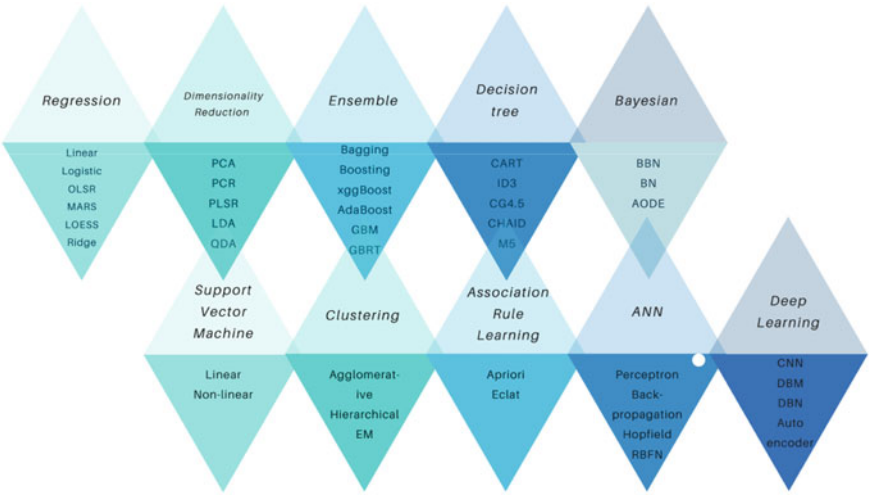


Fig. 59.3 Classification of machine learning algorithms based on similarity

bootstrap sampling method and works by adaptive learning of weaker data instances simultaneously to combine them for maximum accuracy. Random forest is a widely used bagging method that combines randomness and decision trees. Boosting works by learning new hypotheses with each phase of classification sequentially, improving weights of the weaker data instances, and boosts the accuracy. AdaBoost improvises weaker data instances by assigning higher weights to them, and gradient data boosting (GBM) learns with gradients of a loss function [19].

Decision tree, a form of supervised learning, follows a two-step process of growth and pruning of trees. Classification and regression trees (CART) is an algorithm for classifying data into binary labels. CHAID and M5 algorithms are advancements of the CART algorithm. The Iterative Dichotomiser 3 or ID3 algorithm uses information gain to decide the splitting attribute and builds the tree accordingly. C4.5 is an advancement of the ID3 algorithm [20–22].

Bayesian networks (BN) or Bayesian belief networks (BBN) are probabilistic models, and their graphical illustration is in the form of directed acyclic graphs. To overcome the inherent drawback of attribute independence in many real-world settings, averaged one dependence estimators (AODE) was introduced. They are popularly used for statistical validation of any machine learning model [23].

Support vector machines have found applications in real-world applications like face detection, text and hypertext categorization, classification of images, bioinformatics, and generalized predictive control for chaotic dynamics [24].

59.4.2 Unsupervised Learning

The clustering algorithms following under the above category and used widely for implementations include k-means clustering based on centroid approach, mean-shift clustering based on sliding window approach, density-based spatial clustering applications with noise (DBSCAN) based on neighborhood concept in relation to data points. Expectation maximization (EM) and clustering using Gaussian mixture models (GMM) are based on assumption of Gaussian distribution of data points and for finding the parameters of the Gaussian EM. The two variants to hierarchical clustering are agglomerative and divisive hierarchical clustering [25].

Prominently used in market-basket analysis, multimedia data mining, and video data mining, association rule mining has many implementation variants among which, a priori algorithm, frequent pattern (FP) growth remains the most widely used ones [26, 27].

Deep belief networks (DBN) and convolutional neural network (CNN) are extensively used for object detection, whereas the new breed ones yolov3, mask-regional convolutional neural networks (RCNN), and faster RCNN are used in multi-object detection. Auto-encoders are yet another form of neural networks used for dimensionality reduction. Back-propagation remains the backbone of neural network design and mainly used in the design of multi-layer perceptrons [28].

59.5 Market Trends with AI Applications

In this section, we touch upon the broadly used use cases of AI and ML and their diversity in domains. We also look at the region wise distribution of global AI markets. The products dominating the software market have been discussed.

59.5.1 AI Use Cases

AI applications range from healthcare analysis to product recommendations. In a survey by Tractica [29], Fig. 59.4 depicts the top ten predicted to be the highest revenue earning use cases of AI by 2025.

Figure 59.5 [30], on the other hand, illustrates the context for which industries apply ML algorithms and models. Popular reasons why companies incorporate AI include generating hidden insights, improving customer satisfaction, retaining customers by improving interaction, providing prompt query resolution, providing relevant recommendations, improving the customer's overall experience, and thereby reducing churn. Another major reason why companies invest in AI includes reducing costs by forecasting demands and gaining internal insights of the organization.

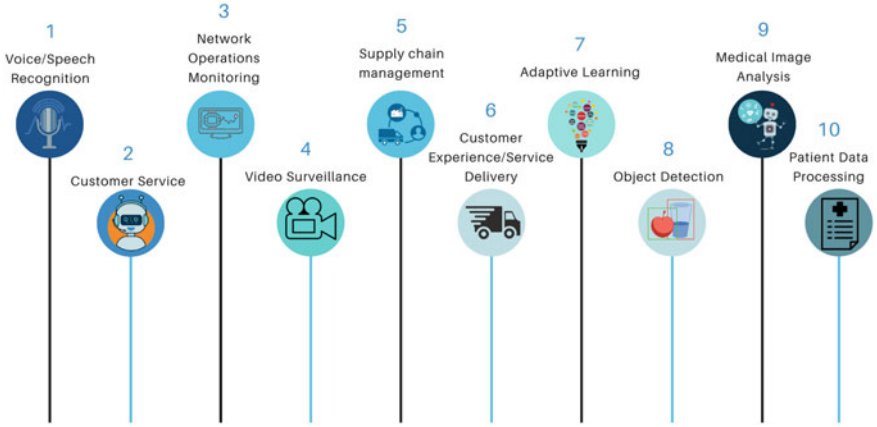


Fig. 59.4 Top ten use cases of AI and ML in software market

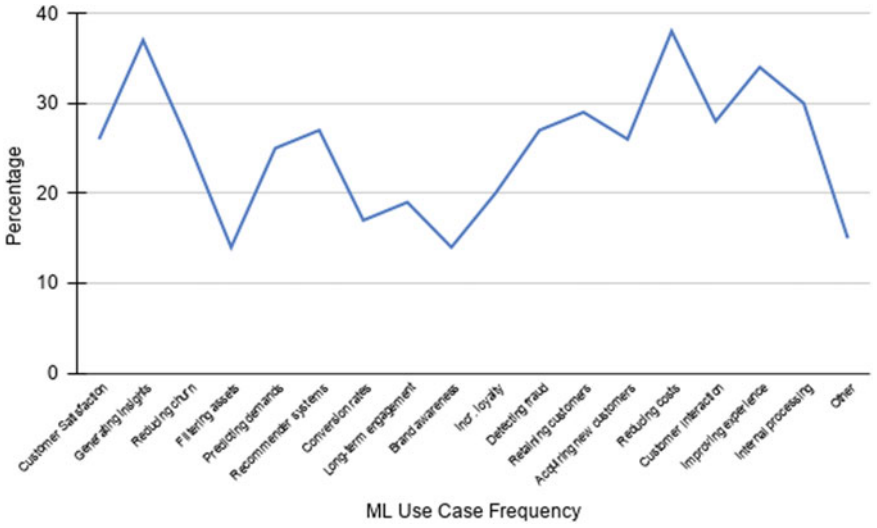


Fig. 59.5 ML use case frequency

Industries also increase loyalty and reduce fraud by imbibing appropriate security mechanisms.

59.5.2 Popular AI-Enabled Products

AI-enabled virtual assistance, delivery robots, and tele-health are some examples of how AI is helping people every day. Use of pseudo-intelligent digital personal assistance (Siri, etc.) to perform a task, ML to accurately detect cancer cells, or Alexa in home automation is definitely going to boost industrial and market demand for AI power products. Google Duplex, Deepmimd's Alpha Star, or AI algorithms providing product suggestions, and support systems in telemedicine are a few applications to which customers have become accustomed to. The personalized ad serving and AI-powered dynamic pricing is nothing less than any boon to various airlines' booking sites. Proliferation in connected devices, data, and usability is trending market growth. Internet dependencies and cloud-based applications in health care, banking financial services and insurance (BFSI), retail, automobiles, security surveillance, smart home devices, and tourism, etc., contain sensitive user information. Data coming from these sources are hence vulnerable to threats and cyber attacks, thus opening a wide scope for designing strong security systems using AI and ML algorithms. Utilizing data from user's wearable devices or smartphones through contact tracing applications for alerting citizens to nearby COVID-19 cases has demonstrated the significance of AI. According to [31], 59% of more people are likely to use tele-health services post-pandemic which was 25% during 2019, while 36% would switch their physician to a virtual one in the USA. This huge demand reports the telemedicine market size to witness $175.5 billion by 2026 from $45.5 billion in 2019 [32]. The wearable AI market alone is forecasted to reach $180 billion by 2025 [33]. In 2019, the global market size for AI by component, technology, deployment, and industry was valued at $27.23 billion and is expected to achieve $266.92 billion by 2027 with 33.2% of compound annual growth rate (CAGR) during this period [34] while another survey [35] calculates a CAGR of 42.2%, with revenue forecast to be $733.7 billion considering solution, technology, and end use.

59.6 Future with AI

Our study so far has highlighted the journey of AI to date. It is a fairly growing market with a lot of future scope. This section discusses the future of AI.

59.6.1 Emerging Trends for the Future

The goal of AI is to design 'general' intelligence for performing unpredictable tasks rather than 'specific' intelligence, which is able to perform only a single task outstripping human skills. The learnings through direct interaction of the object or the model

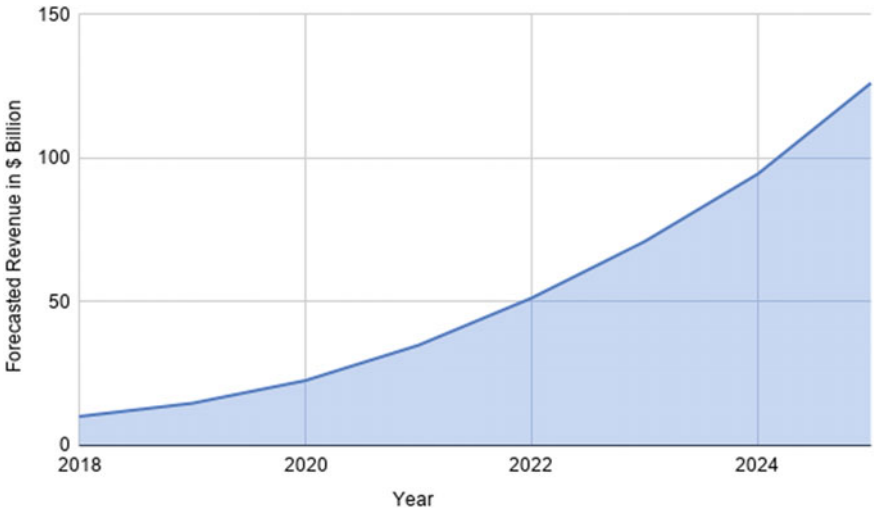


Fig. 59.6 Forecasted market revenue generation from AI software market

to the physical environment are important for domains like robotics and driverless cars for understanding the cause–effect behavior.

The upcoming AI models will be designed based on the pre-trained models and will be implemented on the energy-efficient hardwares [36].

With the vast opportunities and applications AI offers, it is predicted that the global revenue market for AI-driven services and products will drastically increase from 10.1 Billion Dollar in 2018 to 126 Billion Dollars in 2025 as illustrated in Fig. 59.6 [37].

Industries and markets are evolving with business strategies like reducing costs, automation, innovation, and managing risk factors with advanced analytics of AI. This estimates that the markets will grow in the coming few years. As represented in Fig. 59.7 [38], in 2017 in a survey by market research future, it was recorded that 44% of the global machine learning market was centered in North America, 29% in Asia Pacific, 21% in Europe, and 7% for the rest of the world on the basis of the vertical, organization size, and components of the industry.

59.7 Conclusions

Valuable insights into the future market trends and the plethora of opportunities available to the global AI community are focused through the course of this study. Since the nineties when AI first evolved, the field has been growing with dominance. Tools and algorithms are continuously progressing to accelerate hardware performance while optimizing computational resources. In the foreseeable future, AI techniques

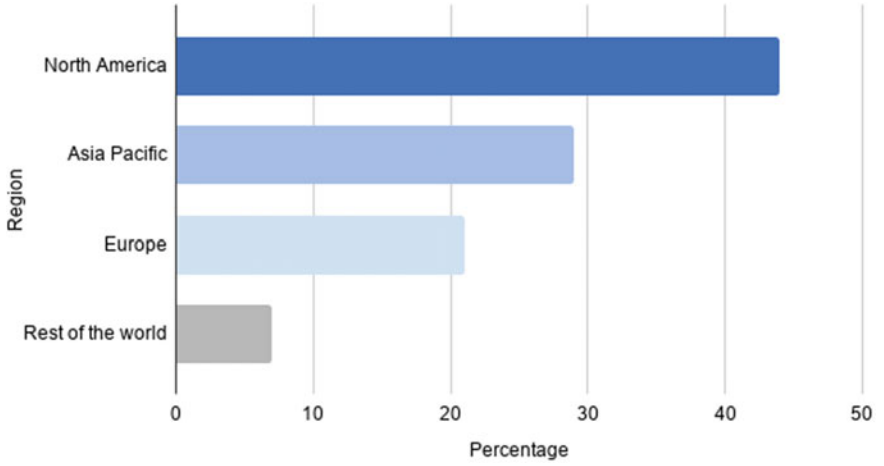


Fig. 59.7 Global machine learning market by region in year 2017

will encompass all unexplored domains such as climate change, clean energy, health care, and education addressing the world's challenging issues.

References

1. Alkrimi, J., Ahmad, A., George, L., Aziz, S.: Review of artificial intelligence. Int. J. Sci. Res. **2**, 487–505 (2013)
2. Oke, S.: A literature review on artificial intelligence. Int. J. Inf. Manage. Sci. **19**, 535–570 (2008)
3. Schmidt, J., Marques, M.R.G., Botti, S., et al.: Recent advances and applications of machine learning in solid-state materials science. NPJ Comput. Mater. **5**, 83 (2019). https://doi.org/10.1038/s41524-019-0221-0
4. Weinreb, D., Moon, D.: The lisp machine manual. ACM SIGART Bull. **78**, 10–10 (1981)
5. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM **9**(1), 36–45 (1966)
6. https://www.forbes.com/sites/gilpress/2016/12/30/a-very-short-history-of-artificial-intelligence-ai/#135cebe96fba
7. Moravec, H.P.: The Stanford cart and the CMU rover. Proc. IEEE **71**(7), 872–884 (1983)
8. Barker, V.E., O'Connor, D.E., Bachant, J., Soloway, E.: Expert systems for configuration at Digital: XCON and beyond. Commun. ACM **32**(3), 298–318 (1989)
9. High, R.: The era of cognitive systems: an inside look at IBM Watson and how it works, pp. 1–16. Redbooks, IBM Corporation (2012)
10. Hassoun, M.H.: Fundamentals of Artificial Neural Networks. MIT Press, Cambridge (1995). https://doi.org/10.1109/JPROC.1996.503146
11. Fogel, D.B.: Evolutionary Computation: Toward a New Philosophy of Machine Intelligence, vol. 1. Wiley, New York (2006)
12. Ryan, C., Collins, J.J., Neill, M.O.: Grammatical evolution: evolving programs for an arbitrary language. In: European Conference on Genetic Programming, pp. 83–96. (1998)
13. Joseph, S., Sedimo, K., Kaniwa, F., Hlomani, H., Letsholo, K.: Natural language processing: A review. Nat. Lang. Process. Rev. **6**, 207–210 (2016)

14. James, G., Witten, D., Hastie, T., Tibshirani, R.: Statistical learning. In: An Introduction to Statistical Learning. Springer Texts in Statistics, vol. 103. Springer, New York (2013). https://doi.org/10.1007/978-1-4614-7138-7_2
15. Nikravesh, M.: Evolution of fuzzy logic: from intelligent systems and computation to human mind. In: Nikravesh, M., Kacprzyk, J., Zadeh, L.A. (eds.) Forging New Frontiers: Fuzzy Pioneers I. Studies in Fuzziness and Soft Computing, vol. 217. Springer, Berlin (2007). https://doi.org/10.1007/978-3-540-73182-5_3
16. Atzori, L., Iera, A., Morabito, G.: The internet of things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010). https://doi.org/10.1016/j.comnet.2010.05.010
17. Draper, N.R., Smith, H.: Applied Regression Analysis, vol. 326. Wiley, New York (1998). https://doi.org/10.1002/9781118625590
18. Vlachos, M.: Dimensionality reduction. In: Sammut, C., Webb, G.I. (eds.) Encyclopedia of Machine Learning and Data Mining. Springer, Boston (2017). https://doi.org/10.1007/978-1-4899-7687-1_71
19. Pandey, M., Taruna, S.: A comparative study of ensemble methods for students' performance modeling. Int. J. Comput. Appl. **103**, 26–32 (2014). https://doi.org/10.5120/18095-9151
20. Safavian, S.R., Landgrebe, D.: A survey of decision tree classifier methodology. IEEE Trans. Syst. Man Cybern. **21**(3), 660–674 (1991). https://doi.org/10.1109/21.97458
21. Anyanwu, M.N., Shiva, S.G.: Comparative analysis of serial decision tree classification algorithms. Int. J. Comput. Sci. Secur. **3**(3), 230–240 (2009)
22. Lasota, T., Sachnowski, P., Trawiński, B.: Comparative analysis of regression tree models for premises valuation using statistica data miner. In: Nguyen, N.T., Kowalczyk, R., Chen, S.M. (eds.) Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems (ICCCI 2009). Lecture Notes in Computer Science, vol. 5796. Springer, Berlin (2009). https://doi.org/10.1007/978-3-642-04441-0_68
23. Neal, R.M.: Bayesian Learning for Neural Networks, vol. 118. Springer Science & Business Media (2012)
24. Noble, W.: What is a support vector machine? Nat. Biotechnol. **24**, 1565–1567 (2006). https://doi.org/10.1038/nbt1206-1565
25. Jain, A.K., Narasimha Murty, M., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. (CSUR) **31**(3), 264–323 (1999)
26. Fang, L., Qizhi, Q.: The study on the application of data mining based on association rules. In: 2012 International Conference on Communication Systems and Network Technologies, Rajkot, pp. 477–480 (2012). https://doi.org/10.1109/CSNT.2012.108
27. Hipp, J., Güntzer, U., Nakhaeizadeh, G.: Algorithms for association rule mining—a general survey and comparison. ACM SIGKDD Explor. Newsl. **2**(1), 58–64 (2000). https://doi.org/10.1145/360402.360421
28. Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M.P., et al.: A survey on deep learning: Algorithms, techniques, and applications. ACM Comput. Surv. (CSUR) **51**(5), 1–36 (2018). https://doi.org/10.1145/3234150
29. Artificial Intelligence Market Forecasts (2020). https://tractica.omdia.com/research/artificial-intelligence-market-forecasts/
30. Algorithma: 2020 state of enterprise machine learning (2020). https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia_2020_State_of_Enterprise_ML
31. Sage Growth Partners, Blackbook Research, 2020. COVID-19 Market Pulse, pp. 1–4. [Online]. Black Book Market Research, United States. Available at: https://blackbookmarketresearch.com
32. Ugalmugale, S., Swain, R.: Telemedicine Market Share Report|Global 2020–2026 Industry Data. [Online]. Global Market Insights, United States (2020). Available at: https://www.gminsights.com/industry-analysis/telemedicine-market
33. Bhutani, A., Wadhwani, P.: Wearable AI market trends—Industry Statistics Report 2025 (2019). https://www.gminsights.com/industry-analysis/wearable-ai-market

34. Fortune Business Insights, 2020. Market Research Report. [Online]. Fortune Business Insights.https://www.fortunebusinessinsights.com/industry-reports/toc/artificial-intelligence-market-100114
35. Grand View Research, 2020. Market Analysis Report. [Online]. Grand View Research. https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market
36. López de Mántaras, R.: The future of AI: toward truly intelligent artificial intelligences. In: Towards a New Enlightenment? A Transcendent Decade. BBVA, Madrid (2018)
37. Revenues from the Artificial Intelligence (AI) software Market worldwide from 2018 to 2025. (2020). https://www.statista.com/statistics/607716/worldwide-artificial-intelligence-market-revenues/
38. Market Key Insight and COVID-19 Impact Analysis (2020). https://www.marketresearchfuture.com/reports/machine-learning-market-2494

Chapter 60
Development of Porting Analyzer to Search Cross-Language Code Clones Using Levenshtein Distance



Sanjay B. Ankali and Latha Parthiban

Abstract Software forking is the process of creating a variant of the existing system by reusing more reliable algorithms and the design of the legacy systems. The software porting team converts the legacy system to the new system by adding or deleting of features and retaining the more reliable algorithm and design. A cross-language clone detector can significantly improve the performance of the software forking team by helping them to locate the code clones of legacy software. This paper proposes a technique to detect functionally equivalent code clones across C, C++, and Java codes by employing the Levenshtein distance algorithm to build a porting analyzer that helps the software forking team with three features like (i) locating all functionally similar cross-language clones among the C, C++, and Java project repositories. (ii) Given an input code for searching, it fetches the functionally equivalent code clone types from the respective repositories. (iii) Displays bug fixes to convert the code from one language to another language. The proposed cross-language clone detector outperforms the accuracy of existing works in terms of precision and can enhance the speed of the software forking process. Results prove that keyword-based filtering using Levenshtein distance detects the cross-language type 1 (Exact) clones and type 2 (Renamed) clones with 100% precision and detects type 3 (Near miss) clones and type 4 (Semantic) clones with 98% and 95% precision, respectively, which is found to be the best accuracy in finding cross-language code clones so far.

S. B. Ankali (✉)
KLE College of Engineering and Technology, Chikodi 591201, India

Research Scholar, VTU-RRC, Karnataka, India

L. Parthiban
Pondicherry University Community College, Pondicherry 605008, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_60

60.1 Introduction

The Internet contains trillion lines of code available nowadays that make the development process much simple [1]. For the same reason, the developers tend not to start coding from the scratch [2]. We find evidence of code reusing in a survey conducted by Vaibhav Saini [3] involving 72 developers, and 96% of the developers prefer searching and reusing code before the start of coding task. The main drawback of using the online source code is maintainability and bug propagation [4] or tends to preach software licenses [5, 6]. Text search engines like Google cannot search the functionally similar code by taking the query as code. Hence, there is a need for a software clone extraction system that takes code as a query and fetches the functionally similar code from the repository by matching the complete code [7]. We propose a technique to accurately classify the cross-language clones of C, CPP, and Java code.

Research Question

1. For the software forking team which is about to convert a large C-based legacy system into C++ or Java-based system containing N separate C code files, is it possible to fetch the corresponding functionally similar codes of C++ or Java from the repository of C++/Java and vice versa with all four clone types classified?
2. Assuming there are three repositories of C, C++, and Java, given a code as a query, is it possible to extract the functionally similar codes of C or C++ with clone types classified into 1, 2, 3, and 4 accurately?

60.2 Background and Related Work

Creating functionally similar programs with or without syntactical change is called software cloning or code cloning [8]. Based on the editing taxonomy, most of the researchers define four basic types of clones [9]. These clone types are grouped into syntactic and semantic classes.

Syntactic: *Type* 1 is also called as exact clones, Type 2 is also called as renamed clones, and Type 3 is near-miss clones. Syntactic clones are based on different editing taxonomies.

Semantic: *Type 4* is semantically similar codes that are different in syntax.

In this section, we briefly define the clone types and present the C and CPP examples to justify our understanding of cross-language clone types which are based on the many great clone detection research in the past [9].

60.2.1 Background

Definition of Cross-language Clones

With the definition of clones mentioned, we can say that two or more codes that produce the same output but implemented in different languages with different four types of variations are called cross-language clones. For an instance, C-code-1 that calculates factorial of a number is called a cross-clone of CPP-code-1 code that calculates factorial of a number [10].

A few example case studies are as follows.

```
1. main()
2. {
3. Int x=100,y=200,z;
4. z= x+y ;
5. printf("sum of two
   numbers=%d", z);
6. }
```

C-code-1

```
1. main()
2. {
3. int x=100,y=200,z;
4. z=x+y;
5. cout<<"sum of two
   numbers is ="<<z;
6. }
```

CPP-code-1

Except for the structure of the language, both the codes above behave similarly. They declare the same two variables x and y, the sum is stored in a variable z, and the result is displayed on the screen. These examples are type 1 cross-clones.

```
1. main()
2. {
3. int a=10 , b=20, c;
4. c = a  +  b;
5.printf("addition program")
6. printf("sum of two
numbers=%d", c);
7. }
```

C-code-2

```
1.main()
2.{
3.int a = 10, b=20,c;
4.c = a + b;
5.cout <<"sum  of  two numbers
  is ="<< c;
6.}
```

CPP-code-2

From the above code snippet, C-code-1 and CPP-code-2 are examples of type 2 clones, whereas C-code-2 is type 3 clone of CPP-code-2.

As the last example, we consider one looping statement.

```
1. int factorial( int n)
2. {
3. int  i, fact=1, n;
4. for ( i = 1;i < = n ; i++)
5. fact = fact * i;
6. return (fact);
7. }
```

C-code-3

```
1. int fact ( int  n)
2. {
3. int i =1,fact = 1,n;
4. while ( i < = n)
5. {
6.   fact = fact * i;
7.   i++;
8. }
9. return(fact);
10. }
```

CPP-code-3

C-code-3 and CPP-code-3 try to find factorial of integer number by using for and while loops, respectively, hence are called type 4 clones. The proposed work will significantly contribute to the software porting process by identifying all the functionally similar code fragments of C, CPP, and Java codes.

60.2.2 Related Work

Past three decades have seen more than 250 tools and techniques in the same language clone detection but very less significant work has happened in cross-language clone detection. In this section, we introduce cross-clone detection with their issues. A tool *CroLSim* [11] is proposed to find software similarity for open-source software clustering. But the results show only 28% precision for searching functionally similar code in a repository. A tool *SLACC* [12] presents the cross-language clone detection based on the input/output behavior of code. The tool lacks three aspects: (i) finds only semantic similarities, (ii) does not classify cross-language clone types (type 1, 2, 3, 4) and (iii) does not support long and complex types of Python D. dead code elimination. A recent study on cross-language source code plagiarism by *Karnalim* [13] works in four steps. But the approach does not handle identifier renaming and (ii) works by converting the original code to intermediate code which is more computationally complex. An AST-based approach proposed by *Perez* [14] is a semi-supervised machine learning model. The system has limitations in (i) matching two codes of Java and Python with only 75% confidence and (ii) does not support more granular clone-type classification (type 1, 2, 3, and 4). A tool *CLCDSA* is proposed by *K. W. Nafi* [11] used action filters to filter out non-probable clones and make the model more scalable. But the approach did not present the results to prove the clone-type classification. A study *Nghi D. Q. Bui* [15] is proposed which is based on bilateral neural networks (BiNN's)-based technique to find similarity and difference in structure based on the language AST using BTBCNN. The method has two issues (i) large programs may slow down the training process and (ii) no results were found to prove the type classification.

Hu et al. [16] proposed the technique using the binary instructions to detect the semantically similar functions through a series of four steps but the approach detects only semantically similar techniques and is vulnerable to the code obfuscation.

An integrated tool *LICCA* [17] is based on the application of the modified longest common subsequence algorithm on enriched concrete syntax tree (eCST) of source code. But the work is limited to semantic clone detection. Clone detection among the.NET language family was proposed by *Al-Omari* et al. [18]. But the work is platform dependent. The work of *Lawton Nichols* [10] is the extension of previous work [19] that finds syntactic similarity using structural and nominal similarity. The method works on the function instead of diffs of VCS and currently handles (C++, Java, and JavaScript). But the system works well only for object-oriented programming languages.

60.3 Proposed Methodology

Figure 60.1 shows the keyword-based four steps to obtain cross-language code clones using the Levenshtein distance algorithm. In step 1, the input code of C, C++, and Java is tokenized. In the second step, it searches all the tokens based on the keywords of the corresponding language and stores the keywords in the single dimensional array.

```
  Algorithm
similarity_Match(String str1,
String str2)
{
String big = str1,small = str2;
if (length(str1) < ength(str2))
    {
       big = str2;
      small = str1;
     }
 int len =length(big);
  if (len == 0)
 { return 1.0;  }
 len= len-edit_Distance(big,
 small)) /(double)len;
       return(len);
 }
```

```
Algorithm
edit_Distance(String s1,
String s2, int len2)
 {
int[] costs = new;
int[length(s2) + 1];
    for i <- 0 to len {
       int old = i;
       for j<- 0 to len {
  if (i == 0)
    costs[j] = j;
     else {
    if (j > 0) {
  int new = costs[j - 1];
 if (s1[i - 1] !=s2[j - 1])
     min((new,old), costs[j])
 + 1);
  costs[j - 1] = old;
  old = new;   }
            }
        }
   if (i > 0)
  costs[length(s2)] = old;
     }
 return costs[length(s2)];
   }
```

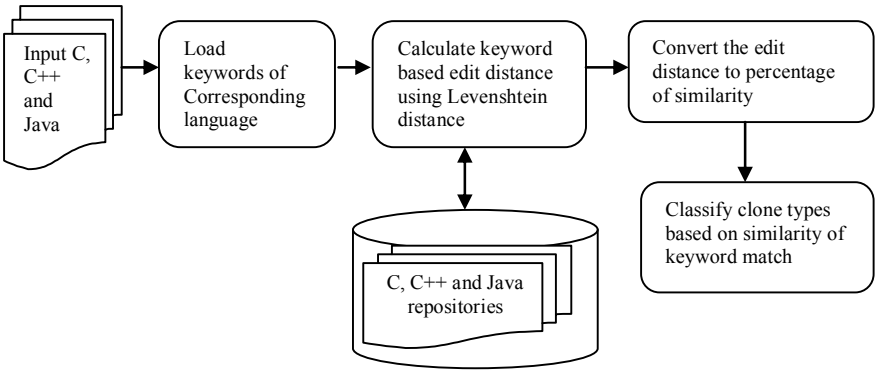


Fig. 60.1 Proposed cross-language clone detection system

Table 60.1 Thresholds of clone-type classification based on similarity values

Similarity matching	Clone type
100	1
≥90	2
≥82 and <90	4
≥68 and <82	3

In step 3, Levenshtein distance is called between the prepared 1D array and all the code documents present in the repositories to calculate the edit distance between the two codes.

As the last step, the classification of clone types is performed by manually validating the matching percentage of all four clone types according to the following thresholds of matching given in Table 60.1.

60.4 Results and Discussion

We calculate the accuracy of the proposed model in terms of precision and recall which tries to find what proportion of positive instances was correct [20].

For instance, we know that C-code-1 and C-code-2 are type 3. If the matching result is type 2, then its true positive results. But if we get the result of matching as any other type, then we call it false positive. With this basic knowledge, we define precision as ***Precision*** = *True Positive/ True Positive* + *False Positive.*

Recall tries to find how many positives are identified correctly by calculating *True Positive/True Positive* + *False Negative*. Figure 60.2 shows the comparison of two project repositories of C and C++ to detect type 1 and type 2 clones between two languages. We have validated the proposed model on the C project containing 136 C codes and 82 functionally similar C++ codes. The framework shows 100% precision in finding the type 1 and 2 clones.

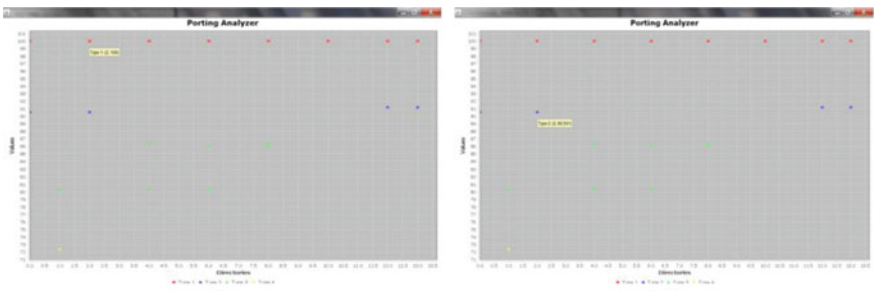


Fig. 60.2 **a** Graph showing common type 1 files between two projects of C and CPP. **b** Graph showing common type 2 files between two projects of C and CPP

Figure 60.3 shows the comparison of two project repositories of C and C++ to detect type 3 and type 4 clones between two languages. We have validated the proposed model with the manual inspection to find type 3 and type 4 clones. The framework shows 98% precision in finding the type 3 and 95% precision in finding four clones.

Figure 60.4 shows the execution of the porting analyzer where the developer tries to search the functionally similar codes of C++ and Java code for the input C code that adds two integer numbers.

Figure 60.5 shows the search results where type 1 code clones of Java and C++ are fetched from the corresponding repositories. Similarly, Fig. 60.4 shows the search results for type 4 code clones of C and Java codes given input as CPP file for finding factorial of a number using do–while loop. The porting analyzer also helps the

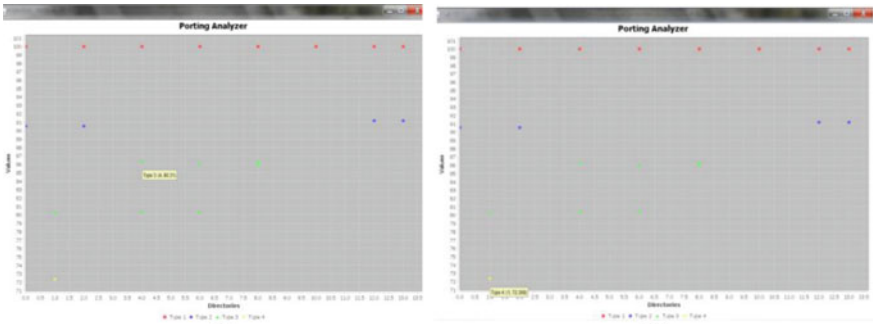


Fig. 60.3 **a** Graph showing common type 3 files between two projects of C and CPP. **b** Graph showing common type 4 files between two projects of C and CPP

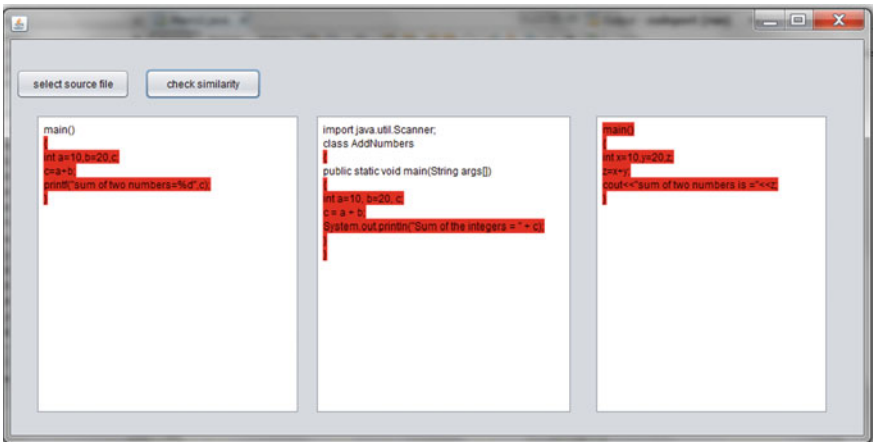


Fig. 60.4 Fetching the similar files of CPP and Java given input as C file for adding two numbers

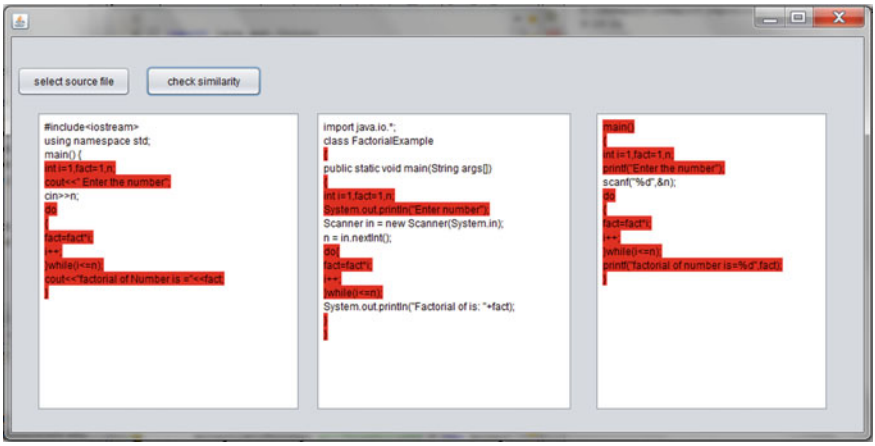


Fig. 60.5 Fetching the similar files of C and Java given input as CPP file for finding factorial of a number using do–while loop

Table 60.2 Performance of our proposed work versus FETT

Technique	Precision (%)	Recall (%)
FETT	2.4	75
Our proposed work	90	100

development team to convert the project from one language to another language by presenting the number of bug fix edits required to convert the functionally similar code from one language to another.

Performance comparison with FETT proposed by Nichols [10].

We have compared the results of our proposed system with the recent work FETT which is freely available in GitHub, and comparison results prove that our system outperforms the FETT in terms of precision and recall which is shown in Table 60.2.

The last feature of proposed porting analyzer is displaying the bug fix while converting the code from one language to another which is shown in Fig. 60.6.

60.5 Conclusion

In this paper, we propose the porting analyzer which is able to detect and locate the cross-language code clones between C, C++, and Java languages with three features like.

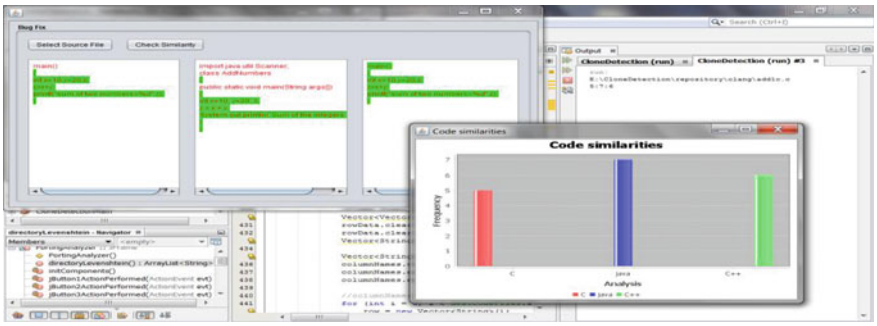


Fig. 60.6 Bug fixes to convert a file from C to C++ and Java

(i) Identifying functionally similar codes across multiple languages (locating C++ and C code given Java code as input and vice versa). (ii) Fetching the complete functionally similar codes across three multiple languages from their respective repositories. (iii) Presenting the bug fixes to convert the code from one language to another across three languages.

The technique can be employed by a software (forking) development team that is involved in converting the legacy software product to a new version of it by addition and deletion of features. Comparison of accuracy in terms of precision and recall with FETT which is only designed for object-oriented programming languages proves that the application of Levenshtein distance using keyword-based searching can be a more accurate way to build the cross-language clone detectors.

References

1. Sadowski, C., Stolee, K.T., Elbaum, S.: How developers search for code: a case study. In: Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering, pp. 191–201 (2015)
2. Acar, Y., Backes, M., Fahl, S., Kim, D., Mazurek, M.L., & Stransky, C.: You get where you're looking for: the impact of information sources on code security. In: 2016 IEEE Symposium on Security and Privacy (SP), pp. 289–305. IEEE (2016)
3. Saini, V., Farmahinifarahani, F., Lu, Y., Baldi, P., Lopes, C.V.: Oreo: Detection of clones in the twilight zone. In: Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 354–365. (2018)
4. Abdalkareem, R., Shihab, E., Rilling, J.: On code reuse from stackoverflow: an exploratory study on android apps. Inf. Softw. Technol. **88**, 148–158 (2017)
5. An, L., Mlouki, O., Khomh, F., Antoniol, G.: Stack overflow: a code laundering platform? In: 2017 IEEE 24th International Conference on Software Analysis, Evolution and Reengineering (SANER), pp. 283–293. IEEE (2017)
6. Baltes, S., Diehl, S.: Usage and attribution of Stack Overflow code snippets in GitHub projects. Empir. Softw. Eng. **24**(3), 1259–1295 (2019)
7. Ragkhitwetsagul, C., Krinke, J.: Siamese: scalable and incremental code clone search via multiple code representations. Empir. Softw. Eng. **24**(4), 2236–2284 (2019)

8. Wang, W., Li, G., Ma, B., Xia, X., Jin, Z.: Detecting code clones with graph neural network and flow-augmented abstract syntax tree. In: 2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER), pp. 261–271. IEEE (2020)
9. Roy, C.K., Cordy, J.R., Koschke, R.: Comparison and evaluation of code clone detection techniques and tools: a qualitative approach. Sci. Comput. Program. **74**(7), 470–495 (2009)
10. Nichols, L., Emre, M., & Hardekopf, B. (2019, April). Structural and nominal cross-language clone detection. In: International Conference on Fundamental Approaches to Software Engineering, pp. 247–263. Springer, Cham
11. Nafi, K.W., Kar, T.S., Roy, B., Roy, C.K., Schneider, K.A.: CLCDSA: cross language code clone detection using syntactical features and API documentation. In: 2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE), pp. 1026–1037. IEEE (2019)
12. Mathew, G., Parnin, C., Stolee, K.T.: SLACC: Simion-based language agnostic code clones (2020). arXiv preprint arXiv:2002.03039
13. Karnalim, O.: TF-IDF inspired detection for cross-language source code plagiarism and collusion. Computer Sci. **21**(1) (2020)
14. Perez, D., Chiba, S.: Cross-language clone detection by learning over abstract syntax trees. In: 2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR), pp. 518–528. IEEE (2019)
15. Bui, N.D., Jiang, L., Yu, Y.: Cross-language learning for program classification using bilateral tree-based convolutional neural networks (2017). arXiv preprint arXiv:1710.06159
16. Hu, Y., Zhang, Y., Li, J., Gu, D.: Binary code clone detection across architectures and compiling configurations. In: 2017 IEEE/ACM 25th International Conference on Program Comprehension (ICPC), pp. 88–98. IEEE (2017)
17. Vislavski, T., Rakić, G., Cardozo, N., Budimac, Z.: LICCA: A tool for cross-language clone detection. In: 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER), pp. 512–516. IEEE (2018)
18. Al-Omari, F., Keivanloo, I., Roy, C.K., Rilling, J.: Detecting clones across microsoft .net programming languages. In: 2012 19th Working Conference on Reverse Engineering, pp. 405–414. IEEE (2012)
19. Cheng, X., Peng, Z., Jiang, L., Zhong, H., Yu, H., Zhao, J.: CLCMiner: detecting cross-language clones without intermediates. IEICE Trans. Inf. Syst. **100**(2), 273–284 (2017)
20. Manning, C.D., Raghavan, P., Schutze, H.: Introduction to information retrieval (2008)

Chapter 61
Multi-Objective Genetic Algorithm for Hyperspectral Image Analysis



T. Subba Reddy, J. Harikiran, and B. Sai Chandana

Abstract The main objective of hyperspectral Image Classification is to group pixels into spectral classes, where each class having a unique label representing specific information in the image. The classification can be done using methods categorized as supervised and unsupervised. The contrast of hyperspectral images is degraded if there is any disturbance of the transmission medium. This disturbance degrades the quality of the image generated by the sensor, which effects the classification accuracy. In this paper, Genetic Algorithm (GA) is used for hyperspectral image analysis. The algorithm is used in contrast enhancement, dimensionality reduction, and classification. This dimensionality reduction will remove less informative bands, decrease storage space, computational load, and communication bandwidth. The experimental results show the improvement of accuracy in classifying Indian Pines and Pavia University datasets

61.1 Introduction

The remote sensing sensors provide data with higher spectral resolution staring from panchromatic to multispectral, multispectral to hyperspectral, and hyperspectral to ultra-spectral. The difference in data collected by the sensor from panchromatic to ultra-spectral is the number of narrow spectral bands and greater spectral resolution. Hyperspectral data contains hundreds of narrow spectral bands. The value of any pixel in a hyperspectral image is not formed by a single value, but by a collection of values in all spectral bands. This collection of values associated with a single pixel is called a spectral signature. This range of values for a single pixel will provide much information about the object in the image.

T. Subba Reddy (✉) · J. Harikiran
School of Computer Science and Engineering, Vellore Institute of Technology, VIT-AP University, Amaravati, Andhra Pradesh, India

B. Sai Chandana
Department of CSE, Gudlavalleru Engineering College, Gudlavalleru, Vijayawada, Andhra Pradesh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_61

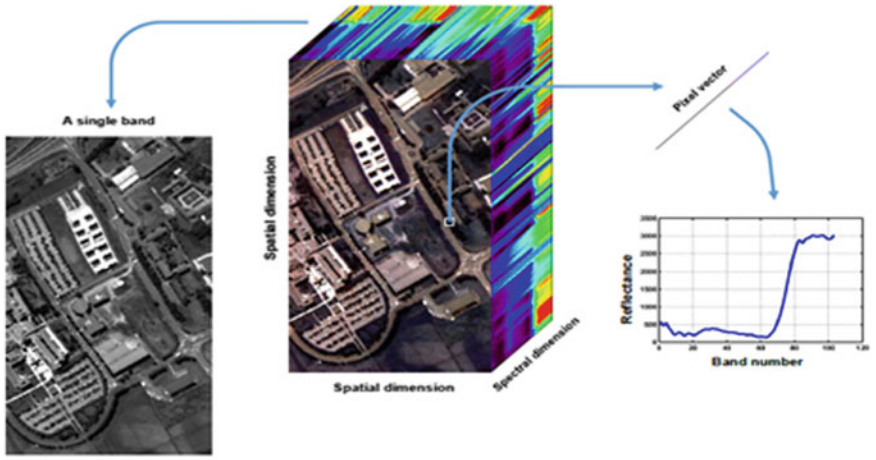


Fig. 61.1 Example of hyperspectral data cube

Hyperspectral Imaging: The sensors in hyperspectral imaging systems are operated from visible wavelength to infrared wavelength, capturing data in hundreds of narrow spectral bands representing the same area of information on the earth. The data for each pixel is not any single value, but it is represented in the form of a vector with each value represents a measurement in that distinct wavelength [1]. The length of the vector is equal to the total number of spectral bands. In hyperspectral data, hundreds of bands are available representing the same scene on earth, while in multispectral data only tens of bands are available. This high dimensionality of data in hyperspectral provides more discriminative information of objects under the scene thus increasing the accuracy of the classification algorithm. Hyperspectral Imaging Systems are used in various applications [2] such as Ecological Science, Geological Science, Minerology, Hydrological Science and Military Applications.

Hyperspectral datasets are called data cubes, where each spectral channel has a specific wavelength representing a grayscale image [3]. Hyperspectral data cube is shown in Fig. 61.1. This paper presents the application of Genetic Algorithm in hyperspectral image enhancement, dimensionality reduction, and classification. The genetic algorithm belongs to a larger class of evolutionary algorithms [4].

61.2 GA Based Hyperspectral Image Contrast Enhancement

Genetic Algorithms [4] is a kind of evolutionary biological process where the solution to any optimization problem is generated using a parallel search in the solution space.

Bi-Dimensional Empirical Mode Decomposition [5]: It decomposes any signal into functions then it is called Intrinsic Mode Function (IMF). This decomposition is a nondestructive process means that summation of these IMFs gives the original signal. The IMFs generated by decomposition process have two specific properties, one is the extreme points and number of zero crossings must be equal or differ by 1, and two the IMFs have zero mean. The first most IMF is a high frequency component and last component which is called as residue is a low frequency component. Summation of IMFs and residue generates the original signal. This signal decomposition using EMD is called shifting process. The same procedure can be extended to two dimensional, which means that decomposition of image named as Bi-dimensional Empirical Mode decomposition [BEMD].

Hyperspectral Image Enhancement Using Genetic Algorithm: The method for image enhancement can be summarized as follows.

i. Given hyperspectral image I is divided into IMFs using 2-D empirical mode decomposition method [5].

$$\mathrm{EMD}(I) = [\mathrm{imf}_1, \mathrm{imf}_2, \ldots, \mathrm{imf}_n] \tag{61.1}$$

ii. The obtained IMFs are used to reconstruct the enhanced image. Each IMF is multiplied with a weight w, and the summation of IMFs with weights gives the enhanced image. The weight for each IMF is obtained using GA.

$$\mathrm{RI} = \sum_{i=1}^{n} w_i * \mathrm{imf}_i \tag{61.2}$$

where w_i denotes the weight corresponding to imf_i, and RI represents the reconstructed image with enhancement.

iii. The GA is used below in enhancement:

(a) First randomly generate the weights of IMFs representing the initial population. The length of elements of each chromosome in the population is equal to the number of IMFs generated by EMD and the element value represents the weight of the corresponding IMF. The values of weights differ between 0 and 1 and the summation of weights is equal to 1.

(b) The information entropy is used as a fitness function for chromosomes in the GA process. The entropy is defined as

$$H = -\mathrm{sum}(p. * \log 2(p)) \tag{61.3}$$

(c) The chromosomes are selected by the Roulette Wheel Technique.

(d) The cross-over rate is equal to 0.7 and the mutation rate equal to 0.07.

(e) The stopping criteria are maximum iterations.

Solution = chromosome with maximum fitness.

The IMF's optimum weights are the values in this chromosome. With these weights and IMFs, the enhanced image is constructed.

61.3 Dimensionality Reduction of Hyperspectral Dataset Using Genetic Algorithm

The huge volume of data, higher dimensionality and redundant information impose negative effect in hyperspectral image analysis. The dimensionality reduction of hyperspectral data was grouped into two types, feature extraction and band selection. The band selection methods are better than feature extraction methods [6], as it selects the effective combination of the original bands by preserving most of the physical characteristics of data without losing any precious details. Some of the existing band selections methods are presented below:

Entropy: The term Entropy measures the data, and it presents in each individual band in terms of uncertainty. The entropy was defined as

$$H(A) = -\sum_{A \in \mathcal{A}} p(A) \log p(A). \tag{61.4}$$

Mutual Information [MI]: [7] The MI is an information measure extended to two variables, one spectral band and another reference image that related to the classification objective. It provides the similarity between two arbitrary variables defined as

$$I(A, B) = \sum_{A \in \mathcal{A}, B \in \mathcal{B}} p(A, B) \log \frac{p(A, B)}{p(A) \cdot p(B)}. \tag{61.5}$$

where $p(A, B)$ defines the joint probability distribution.

Euclidean Distance [ED]: A Vector-Based Distance Measure between vectors X and Y extended to N-dimensions was given by the following equation:

$$\mathrm{ED}(X, Y) = \sqrt{\sum_{k=1}^{N} (X_k - Y)^2} \tag{61.6}$$

The ED measures the vector differences due to magnitude and direction. Band selection was done by choosing the bands having the higher ED values.

Spectral Angle Mapper [SAM] [8]: It measures the Geometric Angle between two vectors X and Y extended to N-dimensions, which corresponds to similarity of shape. Band selection is done by choosing the bands having the higher SAM values. The equation for SAM is as follows

$$\mathrm{SAM}(X, Y) = \arccos\left(\frac{X^T \cdot Y}{\|X\| \cdot \|Y\|}\right) \tag{61.7}$$

Spectral Correlation Mapper [SCM] [9]: It measures the linear relationship between two vectors X and Y extended to N-dimensions is given by

$$\mathrm{SCM}(X, Y) = \frac{\sum_{k=1}^{N_b} (X_k - \mu_X) \cdot (Y_k - \mu_Y)}{(N_b - 1) \cdot \sigma_X \cdot \sigma_Y} \tag{61.8}$$

Band selection is done by choosing the bands having the lower SCM values indicating better separability.

Band Correlation [BC] [10]: It is measure of information redundancy of each spectral band, defined as

$$\mathrm{BC}(i, j) = \frac{\sum_{p=1}^{N_b} \left(x_{ip} - \mu_i\right) \cdot \left(x_{jp} - \mu_j\right)}{\sqrt{\sum_{p=1}^{N_b} \left(x_{ip} - \mu_i\right)^2} \cdot \sqrt{\sum_{p=1}^{N_b} \left(x_{jp} - \mu_j\right)^2}} \tag{61.9}$$

Band selection is done by choosing the bands having the lower BC values, indicating less information redundancy.

GA based Band Selection: The ED, SAM and SCM consider the spectral separability between the bands and BC considers the spectral information between the bands. This new measurement of separability [SEPI] was optimized using Genetic Algorithm (GA) for effective band selection. The SEPI involving ED, SAM, SCM and BC was given by

$$\mathrm{SEPI} = \frac{\mathrm{SCM} \times \mathrm{BC}}{\mathrm{ED} \times \mathrm{SAM}} \tag{61.10}$$

The mechanism of GA based band selection is given below:

(a) Initially we assume population with zeros and ones. Zero denotes bands not selected, one denotes band selected. The Number of selected bands are fixed.
(b) The fitness function for each chromosome

$$\text{fitness function} = \frac{\mathrm{SCM} * \mathrm{BC}}{\mathrm{ED} * \mathrm{SAM}} \tag{61.11}$$

(c) Chromosome selection is done using roulette wheel selection technique.
(d) To reproduce new children GA operators are used to replacing old children.
(e) The number of iterations is fixed to stop the GA Process.

The chromosome with highest fitness value is the solution to the dimensionality reduction. The value 1 in the chromosome denotes corresponding band is selected, and value 0 denotes band is removed.

61.4 GA Based Hyperspectral Image Classification

The Genetic Algorithm as follows:

(a) Chromosomes with K clusters are randomly initialized denoting the initial population.
(b) The multiple objective functions used for calculating fitness of chromosome are K-means index (KMI) [11], Jm measure [12] and Xie-Beni Index (XBI) [13].

$$u_{ik} = \frac{1}{\sum_{j=1}^{K} \left(\frac{D(z_i, x_k)}{D(z_j, x_k)} \right)^{\frac{2}{m-1}}} \tag{61.12}$$

$$z_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m}, \quad 1 \le i \le K \tag{61.13}$$

$$\mathrm{XB} = \frac{\sigma(U, Z; X)}{n \mathrm{sep}(Z)} = \frac{\sum_{i=1}^{K} \left(\sum_{k=1}^{n} u_{ik}^2 D^2(z_i, x_k) \right)}{n \left(\min_{i \neq j} \left\{ D^2(z_i, z_j) \right\} \right)} \tag{61.14}$$

$$J_m = \sum_{j=1}^{n} \sum_{k=1}^{K} u_{kj}^m D^2(x_j, z_k), \quad 1 \le m \le \infty \tag{61.15}$$

$$\mathrm{KMI} = \frac{1}{\sum_{k=1}^{K} \sum_{i=0}^{N} \| x_i - z_k \|^2} \tag{61.16}$$

(c) Chromosome selection is done using Roulette Wheel Technique.
(d) Cross-over rate = 0.8 and mutation rate = 0.07.
(e) Stopping criteria = maximum iterations. Solution = value of chromosome with maximum fitness.

61.5 Experimental Results

The methodology presented in this article is performed on two standard datasets, nine classes Pavia University and sixteen classes Indian pines. 103 bands with each band size 610 × 340 existed in Pavia university dataset. 200 bands with each band size 145 × 145 existed in Indian pines dataset. The datasets are collected from [https://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes]. The image enhancement using Genetic Algorithm is shown in Fig. 61.3. The band selection is done using SPEI measure and the quality of

Table 61.1 Quality of fused image after band selection

Band selection (After fusion using average method)			
Method/India pines dataset	Entropy	Method/Pavia University dataset	Entropy
ED	4.91	ED	4.61
SAM	4.96	SAM	4.78
SCM	4.94	SCM	4.76
BC	4.98	BC	4.81
SEPI	5.09	SEPI	4.89

fused image [using Average Method] by merging the bands selected using each metric was shown in Table 61.1. The SEPI metric selects best bands than other metrics for both datasets. The classification result is shown in Fig. 61.2. By using the methods presented in this paper we get accuracy of classification 92 and 94% for Pavia University and Indian Pines Data.

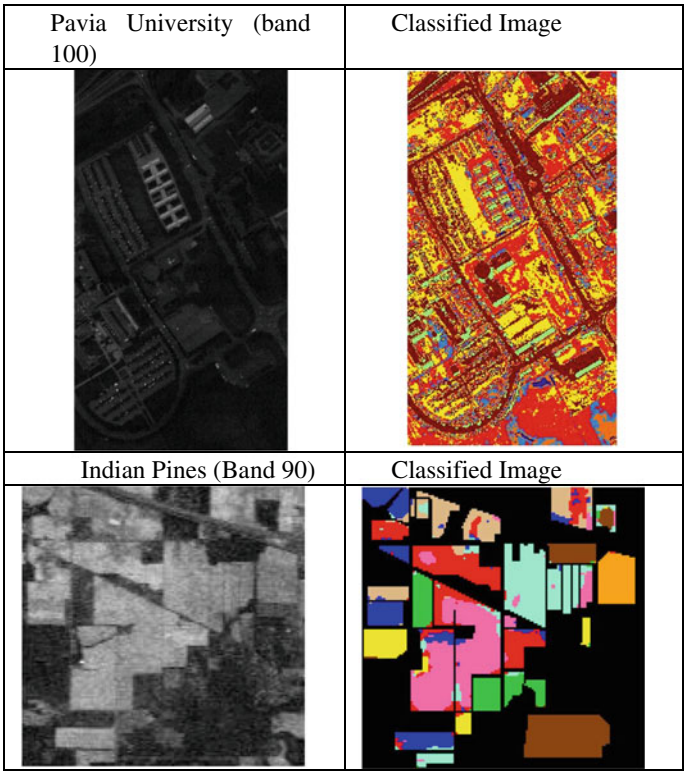


Fig. 61.2 Classification result

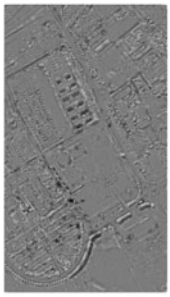
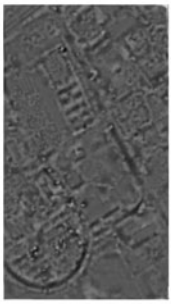
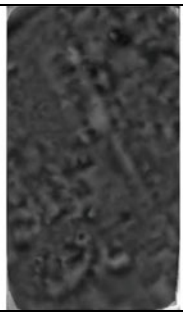
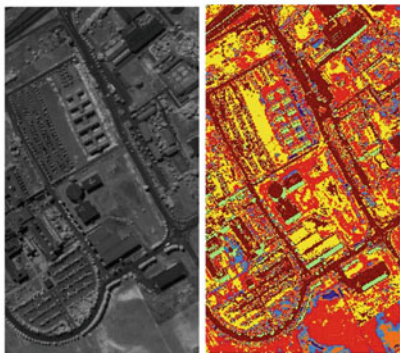
Pavia_University_Image in band_100	IMF-1
	
IMF-2	IMF-3
	
Weights obtained using GA	Enhanced Image Classified Image
Weight of IMF1-0.057 Weight of IMF2-0.103 Weight of IMF3-0.840	
Classification Accuracy (in percentage)	
Before Enhancement: 79% After Enhancement: 92%	

Fig. 61.3 Hyperspectral image enhancement using genetic algorithm

61.6 Conclusions

In this paper, we present hyperspectral image analysis using Genetic Algorithm in enhancement, dimensionality reduction and classification. Instead of optimizing single objectives, multiple objectives are optimized using evolutionary algorithms. In future, many algorithms are used such as ACO, PSO, BCO and Firefox algorithm in a similar way for efficient hyperspectral image analysis.

References

1. Harikiran, J. et.al.: Fast clustering algorithms for segmentation of microarray images. IJSER **5**(10), 569–574 (2014)
2. Saichandana, B. et al.: Image fusion in hyperspectral image classification using genetic algorithm. IJEECS **2**(3), 703–711 (2016)
3. Saichandana, B., et al.: Dimensionality reduction and classification of hyperspectral images using genetic algorithm. IJEECS **3**(3), 503–511 (2016)
4. Holland, J.: Adaptation in natural and artificial systems. Univ. of Michigan Press, Ann Arbor, MI (1975)
5. Zeiler, A. et al.: Empirical mode decomposition—an introduction. In: Proceedings of IEEE IJCNN, pp. 1–8 (2010)
6. Yin, J., et al.: A new dimensionality reduction algorithm for hyperspectral image using evolutionary strategy. IEEE Trans. Industr. Inform. **8**(4), 935–943 (2012)
7. Deb, K. et al.: A computationally efficient evolutionary algorithm for real parameter optimization. IEEE Trans. Evol. Comp. **10**(4), 371–395 (1998)
8. Du, Q. et al.: Similarity based unsupervised band selection for hyperspectral image analysis. IEEE Tran. Geos. RS Lett. **5**(4), 564–568 (2008)
9. Gharaati, E. et al.: A new band selection method for hyperspectral images based on constrained optimization. In: Proceedings of IEEE ICIKT, pp. 1–6 (2015)
10. Hongjun, Su., Qian, Du., Peijun, Du.: Hyperspectral image visualization using band selection. IEEE JSTAEORS **10**(4), 2647–2658 (2013)
11. Tarabalka, Y., et al.: Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. IEEE Trans. Geos. Remote Sens. **47**(5), 2973–2987 (2009)
12. Zhang, H., et al.: Hyperspectral image restoration using low-rank matrix recovery. IEEE Trans. Geo. RS **52**(8), 4729–4743 (2014)
13. Ngatchou, P. et al.: Pareto multi objective optimization. In: Proceedings of 13th ICISAPS. IEEE (2005)

Chapter 62
Auto-Discovery and Monitoring of Network Resources: SNMP-Based Network Mapping and Fault Management



Shaik Imam Saheb and Abdul Rasool Md

Abstract Network management is a service that employs a variety of tools, applications, and devices to assist human network managers in monitoring and maintaining networks. To know the health and state of the network, network operators use network management systems. Network management systems come with fault, performance, configuration management features. Configuration management module helps to keep track of configuration of each device in the network. Often, configuration of the network is manually created. Fault management is the process of locating and correcting network problems or faults. Fault management is an important element of network management which consists of three steps. In the current work, a solution is implemented to discover the network resources automatically and configure them. The system displays a status map, where a visual display of the status of critical network elements is displayed to allow users to verify and isolate problems. An end-to-end testing function is used for testing on a scheduled or on-demand basis. Each network resource deploys an SNMP agent. SNMP agent helps to monitor the state of the network element and notifies the network manager of the events that occur on the element. To notify the network manager SNMP agent uses a collection of objects called MIB. When the MIB objects are received at the network manager, it is parsed for the agent identifying MIB objects, which are displayed on the user interface. The system manager can monitor and modify the MIB objects as per the network user requirements.

62.1 Introduction

Network management refers to the activities, methods, procedures, and tools that pertain to the operation, administration, maintenance, and provisioning of networked systems [1].

S. I. Saheb (✉) · A. Rasool Md
Department of Computer Science & Engineering, Lords Institute of Engineering and Technology, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_62

62.1.1 Network Management Architecture

Most network management architectures [1] use the same basic structure and set of relationships. End stations (managed devices), such as computer systems and other network devices, run software that enables them to send alerts when they recognize problems (e.g., when one or more user-determined thresholds are exceeded). Upon receiving these alerts, management entities are programmed to react by executing one, several, or a group of actions, including operator notification, event logging, system shutdown, and automatic attempts at system repair.

Management entities also can poll end stations to check the values of certain variables. Polling [1] can be automatic or user initiated, but agents in the managed devices respond to all polls. Agents are software modules that first compile information about the managed devices in which they reside, then store this information in a management database, and finally provide it (proactively or reactively) to management entities within network management systems (NMS) via a network management protocol [2] as shown in Fig. 62.1. Well-known network management protocols include the simple network management protocol (SNMP) and common management information protocol [3] (CMIP).

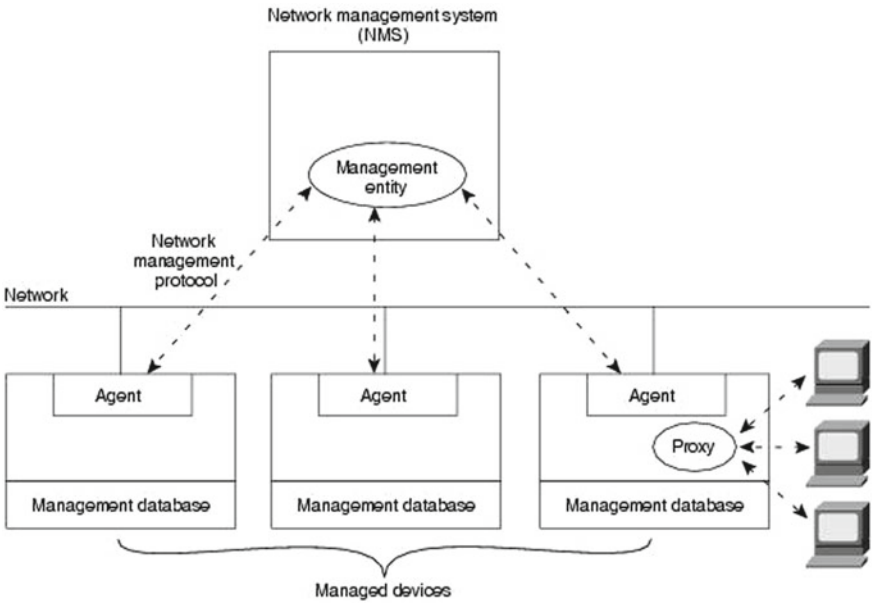


Fig. 62.1 SNMP network management system

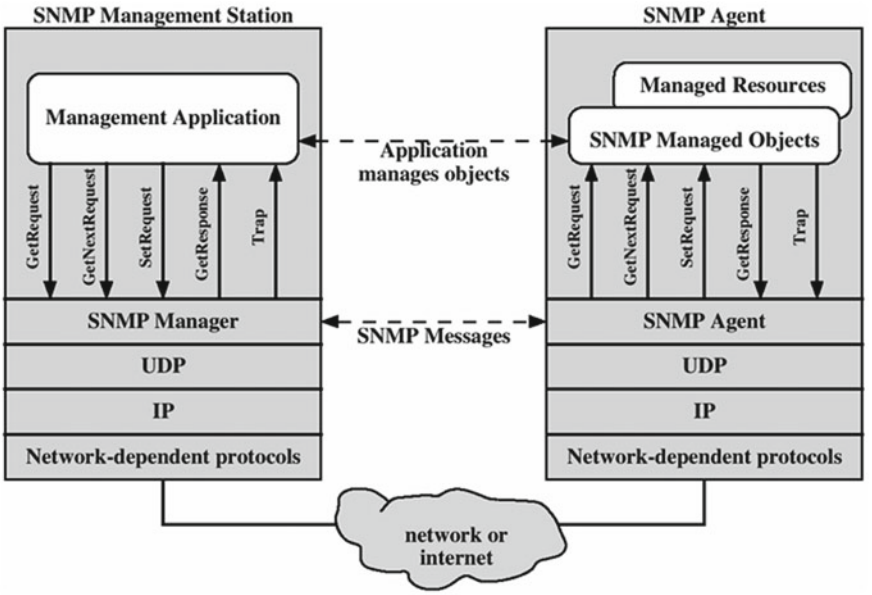


Fig. 62.2 Communication between SNMP management station and SNMP agent

62.2 Related Work

Configuration and monitoring of network elements is essential in communication networks to ensure that the network elements are up and running. To monitor and configure, we need to have a view of the network. This chapter discusses various protocols, techniques, methods of ICMP and SNMP that are helpful to develop this paper.

62.2.1 Internet Control Message Protocol (ICMP)

Internet Control Message Protocol (ICMP) is used to communicate specific information between hosts about the network communications problems. The ICMP consists of different messages in order to communicate among its hosts.

62.2.1.1 ICMP Message Types

To communicate among the host systems, computers exchange messages. The message which are exchanged by ICMP is defined as

Table 62.1 Components are the building blocks of SNMP

1	SNMP management systems and agents
2	Management information base (MIB)
3	SNMP messages
4	SNMP communities
5	The communication process between SNMP managers and agents

62.3 SNMP Architecture

SNMP is based on the manager/agent model consisting of a manager, an agent, and a database of management information, managed objects and the network protocol. The manager provides the interface between the human network manager and the management system. The agent provides the interface between the manager and the physical device(s) being managed, such as bridges, hubs, routers, or network servers, and these managed objects might be hardware, configuration parameters, performance statistics, and so on.

62.3.1 SNMP Agents and Managers

There exist two types of entities in SNMP environment, managers and agents. A manager, or network management station, is a device or software that is managing one or more agents. A network management station is used for fetching data from agents. It receives and handles traps and responses from agents. An agent, on the other hand, is operating on the managed device. The agent is implemented to send traps on different events. It also responds to requests from a manager as shown in Fig. 62.2. The implementation of an agent is based upon MIB files. The basic components of SNMP are listed in Table 62.1.

62.3.2 SMNP Data Structures

The data handled by the SNMP agent is organized into units that are called "management information bases" or MIBs for short. MIBs are described via a precise definition language called "Abstract Syntax Notation" [4]. A manager can talk to an agent, understand the agent information, and handle it if the manager has the MIB of the agent.

62.3.2.1 The Management Information Base

The manager and agent use a management information base and a relatively small set of commands to exchange information. The MIB is organized in a tree structure with individual variables, such as point status or description, being represented as leaves on the branches.

62.4 Design and Implementation

62.4.1 Auto-Discovery and Network Monitoring Using ICMP and SNMP

A networked application sends ICMP echo requests on to the subnet to find the network addresses to be discovered. Responses are parsed to identify the IP addresses of the machines, and SNMP get requests are used to know the gateways, routers in the network. This solution requires ICMP echo messages to be sent to each device in the network and get the IP address as the response (seed address) to perform the initial discovery process. With the initial seed address, we do the broadcasting or echoing to get all other devices specified with in the range. The response messages are being parsed into message and are sent to the NetVision application to update/show the available network IPs on to the user interface. Polling is done every 4 min to know the status of devices in the network. This application enables us to know the IP address in the network.

The gateways, routers, or switches have a feature called ipForwarding. An SNMP get request is sent to the devices on the network for the OID ipForwarding of MIB-II of RFC-1213. If the response for the get request is TRUE, it means the device is an IP forwarding device. If ipForwarding is TRUE, a request is sent to the device to identify the number of interfaces on the device by sending an SNMP get request on ifNumber of MIB-II. If the device has multiple interfaces, then it may be identified as a router.

Discovery of hosts:

Once router is identified, a SNMP request is sent to it on the OID ipNetToMediaTable using interface-index value from ipAddrTable and ipNetToMediaTable, the subnet and host IP of each device is found. The IpAddresses are given as input to fetch and set attributes values for the MIBs system, interface, IP and host resource.

a. **Implementation of SNMP Agent**
   Init_agent().
   This routine initializes the agent. The exact initialization is

1. Creation of sockets to receive SNMP requests
2. Initialization of logging mechanism

3. Installation of signal handlers
   Init_mib_modules().
   The MIBs to be deployed in the SNMP agent at run time environment are dynamically loaded into the program memory by using this routine. system_mib, ip_mib, interfaces_mib and hr_mib are loaded into agent run time environment by invoking the init routines of the respective MIBs.
   Receive().
   This is the infinite loop for the SNMP agent daemon application.
   Exit criteria are A signal is sent by the user to shutdown this application.
   Select().
   This system call helps to listen on the sockets in the blocking and non-blocking modes. Current implementation of agent environment has no job than listening on the socket. Hence, select system call runs in the blocking mode.
   Receivefrom().
   This system call is used to receive an incoming message from the socket. Receivefrom is applicable in case of UDP. As client and server run over UDP transport to receive messages the system call used is receivefrom.
   After all the variables in the variable list are acted upon a response PDU is constructed and sent to the source of the request message (Note: reply is encoded before it is sent to the source).

b. Implementation System of MIB
   System_variables is an array of structures with each node having the following information magic number to index into the structure array, its ASN data type, level of access (read only or read–write), pointer function that accepts requests on this MIB, depth of the variable in the MIB tree, and node ID of the object under the node system in the MIB tree.

```
void init_system_mib(void);
```

MIB initialization is done by this routine. This routine is invoked by agent at the time of agent initialization. REGISTER_MIB is the routine which registers the MIB with the agent.

```
REGISTER_MIB(mibII/system, system_variables, variable2,
system_base_oid);
```

var_system is invoked by the find_var_method (the pointer function invoked by handle_one_var in the agent) when the incoming request is meant for system_mib. Irrespective of whether the request is GET or SET var method is invoked. If the request is GET, return value of this routine is used to populate response message. If the request is SET, write method (in this context writeSystem) is invoked after var method returns (Fig. 62.3).

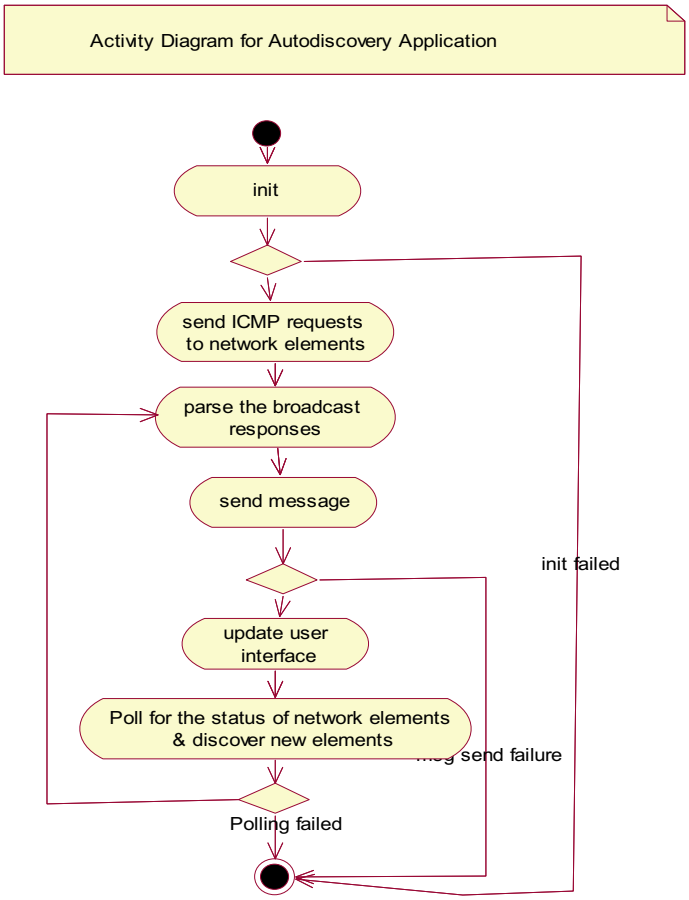


Fig. 62.3 Activity diagram for auto discovery application

62.5 Discussion

See Figs. 62.4, 62.5, 62.6 and 62.7.

62.6 Conclusion

The primary objective of this paper was to discover network resources automatically. This application writes a configuration file with the discovered network resource information. SNMP MIBs are implemented for system-MIB, interface-MIB, IP-MIB and host resource-MIB variables. Variables are retrieved and set on the MIBs.

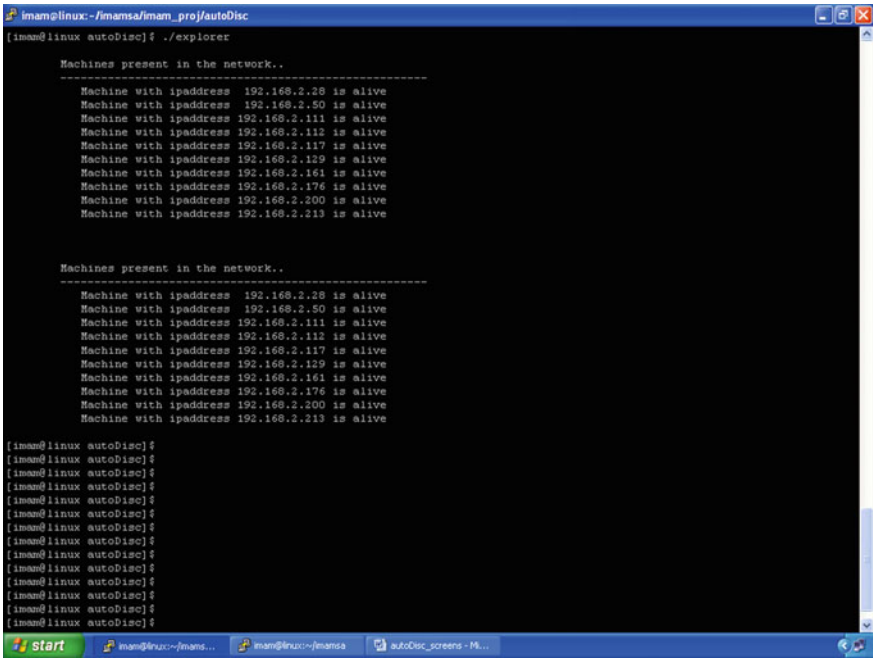


Fig. 62.4 Polling is done for every four minutes to find new network elements which are added or the existing elements that are being removed if the system goes down

The combination of ICMP and SNMP protocol drastically reduces the time taken to discover the network resources and monitor their status.

Future Work.

The future work related to this paper could focus on auto-discovery of network topologies and display the resources in a graphical format. Some of the further investigations are listed below.

- SNMP client applications can be integrated with auto-discovery application to provide a graphical user interface and to support configuration and fault management features of network management.
- Auto-discovery solution can be augmented to support expansive network topologies (WAN).
- MIBs can be implemented to monitor, check the status and alarm various computer devices as well as non-computer devices (like mechanical, electrical, robots, infrastructural works) when they meet their threshold values.

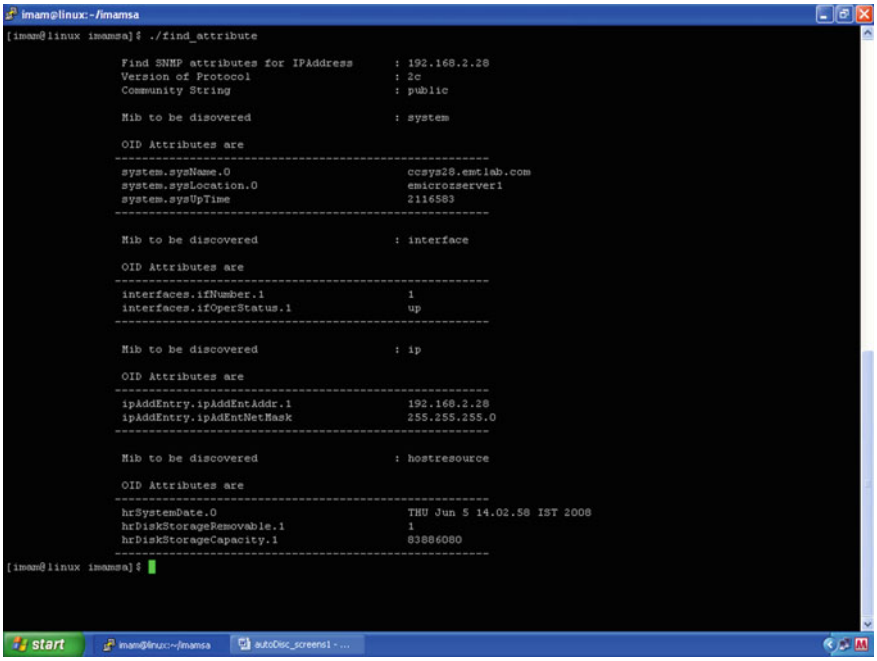


Fig. 62.5 When find_attribute is launched for the network IPAddress 192.168.2.28 its respective MIB objects(attributes) are retrieved

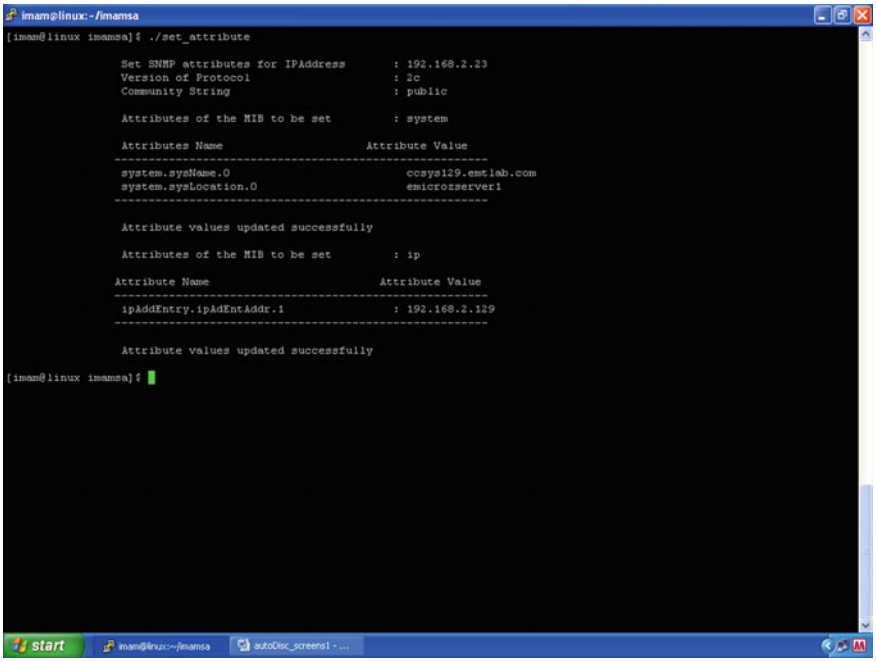


Fig. 62.6 set_attribute application is launched from command line. It asks for the IP address of the network element to be set. Upon giving the IP address, it asks for the version, community of SNMP protocol to be used. In the implementation, currently SNMPv2 is used. MIB object (variable) to be set is to be given as input. Upon giving valid input, it forms a command message and invokes snmpset request, and the response returned from snmpset is shown on the screen. The IPAddress provided as input is 192.168.2.23

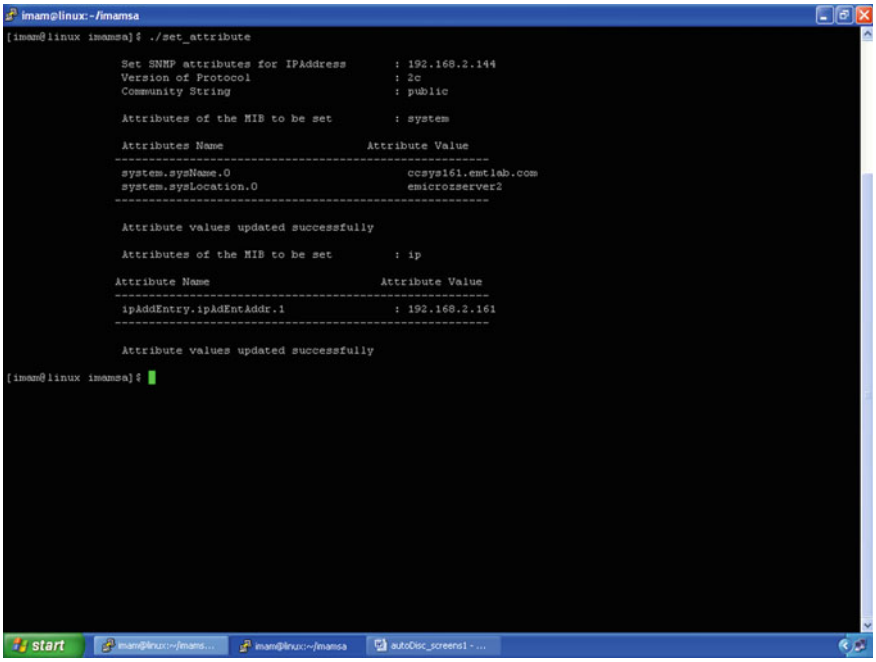


Fig. 62.7 The MIB objects of the IPAddress 192.168.2.144 are being set or initialized to new values

References

1. IETF: Configuration management with SNMP (2000). https://www.ietf.org/
2. Case, J., McCloghire, K., Rose, M., Waldbusser, S.: Protocol Operations for Version 2 of the Simple Network Management Protocol (SNMPV2), RFC 1905 (1996)
3. Perkins, D.T.: Snmp versions. The Simple Times, **5**(1), 13–14 (1997)
4. IETF: The Internet Engineering Taskforce (2000). https://www.ietf.org

Chapter 63
Privacy Preserving Semantic-Aware Multi-keyword Ranked Dynamic Search Over Encrypted Data



M. D. Asrar Ahmed, S. Ramachandram, and Khaleel Ur Rahman Khan

Abstract Searchable encryption (SE) is a promising primitive which enables data owners to privately query and retrieve required documents from encrypted outsourced data collection in cloud servers. Most of the SE schemes in literature strive to achieve a balanced trade-off between efficiency, privacy and expressiveness of query. However, a vast majority of efficient and secure schemes lack searching capabilities beyond exact match of queries and ignore the semantic relationship between keywords and documents. Therefore, a new trend in SE research is to address the question: How to make search semantic-aware and context-based? Majority of the existing semantic search schemes are static, do not support dynamic update operations, suffer from leakage of significant information to server, and employ huge size vectors to represent documents, keywords and topics/features. We propose a privacy preserving, relevance ranking-based dynamic search scheme using a sublinear search time hash-based indexing mechanism and semantic-aware search supporting semantic expansion of user queries. Our scheme employs Word2Vec model to quantitatively evaluate semantic relationships between query and other keywords to expand queries to fulfill user's search intentions. It also supports dynamic operations like insert and delete of documents on outsourced collection with less efforts to update relevance information which is calculated partially at client and server. Experiments on RFC dataset have demonstrated a time and space efficient non-interactive semantic search over encrypted data.

63.1 Introduction

A potential bottleneck for cloud users to outsource huge volumes of confidential data despite associated benefits is the fact that cloud service provider is not fully trusted.

M. D. Asrar Ahmed (✉) · S. Ramachandram
CSE Department, University College of Engineering, Osmania University, Hyderabad, India

K. U. R. Khan
CSE Department, Ace Engineering College, Rangareddy, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_63

Searchable encryption (SE) is a solution to this problem, which enables data owners to privately search the outsourced encrypted dataset. With first practical SE scheme proposed in year 2000 by Song et al. in [1], many researchers have contributed to the field of SE. The subsequent research works include: single keyword search [2], multi-keyword search [3], and fuzzy search [4].

The major drawback of most of these schemes is that they consider only exact match of the keywords, and very few consider fuzzy searches. And use the very common TF-IDF model to draw relevance between document and keyword. In general user intention during search is mostly semantic in nature than being syntactic. For example, a user searching for "economy" keyword in document collection should not be restricted to just "economy", but the user intention will be to retrieve documents which contain keywords semantically closer to it, such as "finance" and "monetary". The TF-IDF model does not take into consideration such semantic relationships/associations, as it ranks keywords based on mere occurrence and more so syntactically and it also requires dealing with high-dimensional and sparse vectors representing documents and keywords due to huge size of keyword dictionary. Additionally, most of the schemes use hierarchical index structures, populated with prior ranking information, causing dynamic updates at several levels in it, to be a costly affair.

In this paper, we propose a semantic-aware multi-keyword dynamic search scheme over encrypted data returning top-k ranked documents. We achieve semantic expansion of queried keywords using the Word2Vec model [5] to meet users' search intentions. The matching documents are returned to user based on TF-IDF-based ranking of documents. The subsequent dynamic operations such as insert/delete operation on documents are carried out without affecting other unrelated index entries as document ranking is computed partially at client and server. Our proposed scheme uses hashing-based indexing mechanism by Etemad et al. [6], with significant optimizations with respect to support for dynamic operations, semantic search, and improved security.

63.2 Related Work

The searchable encryption (SE) schemes have evolved since the pioneering work proposed by Song et al. [1] which was first practical solution with linear search time. It was followed by many schemes with improved efficiency (sublinear, logarithmic in size of document collection) and supported expressive queries. The first Semantic Search scheme over encrypted data was proposed in [7] makes use of stemming algorithms to obtain root of a query keyword and does the search. In [8], authors proposed a semantic search scheme which obtains similar keywords to query with help of edit distance as quantitative similarity measure. In [9], authors proposed a synonym query scheme, where a common synonym thesaurus was built on foundations of NARCT. Sun et al., in [10] proposed a semantic search scheme using two separate servers, expanding the query using semantic relationship library employing

Table 63.1 Comparison of existing strategies with proposed work

Scheme	Search complexity	DSSE	Index type	Semantic search	Forward privacy	Multi-keyword
[5]	$O(\|S\|)$	No	Inverted index	Stemming	No	No
[6]	$O(\tau \|W\|)$	No	Tries	Edit distance	No	No
[7]	$O(a_w)$	No	Vector space	TF-IDF	No	Yes
[8]	$O(\|S\|)$	No	Inverted index	MI	No	No
[9]	$O(\theta \|W\| \log n)$	No	Secure k-NN	SRG	No	Yes
[10]	$O(mn^2)$	No	Secure k-NN	ECH	No	No
[11]	$O(mn)$	No	Secure k-NN	LDA	No	Yes
[12]	$O(n)$	No	Secure k-NN	Word2Vec	No	Yes
[13]	$O(n)$	Yes	Vector space	Doc2Vec	No	Yes
Proposed	$O(s_w)$	Yes	Hash-based index	Word2Vec	Yes	Yes

Notations |S| is size of stem set,τ : size of similarity set, $|W|$, m: dictionary size, a_w: # documents matching keyword, SRG: semantic relationship graph, θ: # of leaf nodes, ECH: extended concept hierarchy, $|D|$, n: # of documents, LDA: latent Dirichlet allocation, s_w : # of semantically related keywords to w

Mutual Information metric defined in [11]. In [12], authors built Semantic Relationship Graph (SRG) by computing Mutual Information. In [13], authors used Extended Concept Hierarchy with help of WordNet tool. Dai et al., in [14], proposed a scheme based on Latent Dirichlet Allocation topic model by generating keyword and document relevance matrices for efficient similarity computations. Hsieh et al. in [15] have proposed a semantic search scheme using Word2Vec machine learning-based model. The work in [16] is based on Doc2Vec model and uses secure k-NN algorithm to compute similarity between documents and keywords. A detailed comparison of these schemes with our proposed scheme is given in Table 63.1.

63.3 Proposed Scheme: Notations, Semantic Search, and Algorithms

63.3.1 Notations

- Documents collection: $D = \{d_1, d_2, \ldots, d_n\}$, encrypted documents $C = Enc(F)$
- K_1, K_2, K_3: keys generated by $KeyGen(1^\lambda)$ used in $Enc_{K_1}()$, $P_{K_2}()$, $G_{K_3}()$
- $DictW$, $DictD$, $DictV$: Keyword, Document, and Vector dictionaries

- $f_{d_{j,t}}$: term frequency of word $w_t \in W$ in file $d_j \in D$, f_{w_t}: # files containing term w_t
- V_t: m dimensional vector for word w_t generated by Word2Vec model.
- $Enc_{K_1}()$, $P_{K_2}()$, $G_{K_3}()$: Symmetric Encryption algorithm, a Pseudorandom Permutation, and a Pseudorandom function family, respectively.
- $I = \{(i, w_i)|w_i \in W\}$, be an index of all unique keywords in set W.
- $T_w =< P_{k_2}(i), x_i, f_{w_i} >$, a hash table at client side (client state).
- $\forall\, w_i \in d_j, \forall 1 \leq i \leq m and \forall 1 \leq j \leq n$ compute: $tf_{ij} = \frac{1}{|d_j|}.(1 + \ln f_{d_j,i})$.

63.3.2 Semantic Extension of Queries—Word2Vec Model

In order to match the users' search intentions in addition to exact match of query keywords, semantic search extends the query terms based on semantic relationship with other terms. Quantitative evaluation of semantic relationship between keywords is achieved by employing the machine learning-based Word2Vec model [5]. Word2Vec as shown in Fig. 63.1, works by generating vectors v_t of some predefined dimensions representing each word $w_t \in W$, and mapping it to a window of surrounding context words, using a neural network approach. And the semantic similarity between any two words w_q and w_j is computed with respect to cosine similarity of their vectors V_q and V_j as follows (Eq. 63.1):

$$CS(V_q, V_j) = \frac{V_q.V_j}{\|V_q\| \cdot \|V_j\|} \tag{63.1}$$

$$w_k = \text{argmax}_{w \in W}\{CS(V_q, V_k)\} \tag{63.2}$$

The keyword w_k which is semantically closest to query keyword w_q is computed using Eq. 63.2. We store these vectors in the Word Vector Dictionary $DictV$ based on unique index assigned to keywords in index set I. Therefore, the vectorization of keywords is oblivious to the server.

63.3.3 Proposed Scheme

We present below a privacy preserving semantic-aware multi-keyword dynamic search scheme by adopting the forward secure dynamic searchable encryption scheme proposed in [6] and achieve the semantic search by employing the Word2Vec model.

Setup and Index Building: The different secret keys used during index building and searching are generated using a security parameter λ and a probabilistic key

generation function: $KeyGen(1^{\lambda})$. The vectors set $V = \{V_i | \forall i \in [m]\}$, is generated to quantitatively summarize the keywords using Word2Vec model specifying a dimension of size l. This set is outsourced to server with each vector linked to keyword specific key K_{w_i}. The $DictW$ dictionary stores label-value pairs of the form: (hash address, randomized term frequency value along with document identifier) as described in $BuidlIndex$ algorithm. The $DictD$ dictionary is used to store hash address of $DictW$ entries to act as forward index used during updates. The data owner maintains a local hash table T_W as its state consisting of entries: a random integer x_i for each keyword w_i and the term frequency f_{w_i}. A keyword specific key K_{w_i} is generated for each keyword through $G_k()$ with keyword and x_i, which changes with every search/update operation. This x_i is used to ensure randomized queries to prevent keyword guessing attack as well file injection attacks [17], and also, the previous queries cannot be re-run due to ever changing random input to trapdoor generation algorithm. This specific key is then used as an input to the keyed hash function h_1 along with keyword-document counter to generate a hash address to store/retrieve entries.

Algorithm 1: *BuildIndex*

Input: $W, D, I, K_2, K_3, PRF\ \ G$:

Client Steps:

1. Set $T_w =< P_{k_2}(i), x_i, f_{w_i} >$, initially x_i=0 $\forall\ w_i \in W$, of size |W|. $f_{w_i} = 0$
2. $\forall w_i \in d_j, \forall 1 \le i \le m\ and\ \forall 1 \le j \le n$: $tf_{ij} = \frac{1}{|d_j|} \cdot (1 + \ln f_{d_j,i})$
3. Build the vectors for all $w_i \in W$, using Word2Vec model: V_i of dimension l.
4. $for\ all\ d_j \in F$:
   a. $generate\ a\ key\ for\ d_j$: $K_j = G(K_3, id(d_j)||R_2)$, where R_2 is a random number.
   b. $for\ all\ w_i \in d_j$: //$1 \le i \le |d_j|$
      i. $dlabel_i = H_2(K_j, i)$
      ii. $if\ T_w[P_{k_2}(i)].x_i = 0\ then\ x_i = randInt()$
      iii. $T_w[P_{k_2}(i)].f_{w_i} = T_w[P_{k_2}(i)].f_{w_i} + 1$
      iv. $K_{w_i} = G(K_3, w_i||x_i)$
      v. $label_i = H(K_{w_i}, T_w[P_{k_2}(i)].f_{w_i}||0)$
      vi. $dval_i = \langle label_i \rangle$
      vii. $val_i = (\langle id_j, tf_{ij} * R_1 \rangle \oplus H(K_{w_i}, T_w[P_{k_2}(i)].f_{w_i}||1), dlabel_i)$
      viii. $AllPairsD = \{dlable_i, dval_i\} \cup AllPairsD$
      ix. $AllPairsW = \{(label_i, val_i)\} \cup AllPairs$
      x. $AllPairsV = \{\langle V_i, K_{w_i} \rangle\} \cup AllPairsV$
   c. Encrypt d_j: $c_j \leftarrow Enc_{K_1}(d_j)$
5. $C = \{c_1, c_2, \ldots, c_n\}$, $AllPairsW$, and $AllPairsD$ is sent to Server

Server Steps: Creates three empty Maps $DictW$, $DictD$ and $DictV$ to store values from sets $AllPairsW$, $AllPairsD$ and $AllPairsV$ respectively.

Search Operation: The search operation begins with data user generating a trapdoor for query $Q = \{w_1, w_2, \ldots, w_q\}$ by semantically expanding Q. The data user after appropriate authorization from data owner, generates following inputs to send to server: the existing keyword specific keys each of w_q (used to locate entries in $DictW$) the new keyword specific keys for each of w_q (to re-insert entries in

$DictW$), file count of each of queried keywords, i.e.,$T_w[w_q].f_{w_q}$. On receiving trapdoor$\langle K, nK, F\rangle$, server parses it to semantically extend the query by computing the similarity score through Eqs. 63.1 and 63.2 using Word2Vec model. The top similarity scores are considered based on certain threshold value to enrich the query with respect to semantics. Now the semantically extended query $Q' = \{w_1, w_2, \ldots, w_{q'}\}$ where$q' > q$, along with keys associated with newly added keywords to query from $DcitV$ are used to retrieve document identifiers from $DictW$ by computing relevance score of each document to extended query keyword. The server runs the search algorithm for every queried keyword as described in step 4 of $SemanticSearch$ algorithm. The rank of a document with respect to semantic query keywords is computed using randomized term frequency value contained in above format of entry astf_{ij},n, and f_{w_i} as in step 4. The sums of scores of each document for all queried keywords are calculated, and finally, those documents which contain all queried keywords are sorted in ascending order based on their ranking score and these documents are then sent in response to user query.

Algorithm 2: *SemanticSearch*

Client Input: Old and New Keyword-specific keys $K = \{K_{w_1}, K_{w_2}, \ldots, K_{w_q}\}$, $nK = \{K_{w_1}, K_{w_2}, \ldots, K_{w_q}\}$, word-file counts $F = \{f_{w_1}, f_{w_2}, \ldots, f_{w_q}\}$ from the T_w table.
Server Input: $DictW, DictD, DictV$, and n
Client Steps:

1. $Q = \{w_1, w_2, .., w_q\}$, with K and nK generated as follows:
   a. $for\ all\ w_q \in Q$
      i. Generate random integer for w_q as x_q in $T_w[w_i]$.
      ii. $K_{w_q} = G(K_3, w_q||x_q)$, $nK = nK \cup \{K_{w_q}\}$
      iii. $F = F \cup \{T_w[P_{K_2}(q)].f_{w_q}\}$
   b. Send $\langle K, nK, F\rangle$ as a conjunctive keywords query to server.

Server Step:

2. Extend the query using Word2Vec model and computer similarity scores using equation (1 & 2) to obtain new token: $\langle K', nK, F'\rangle$
3. Initialize a $ResultSet = \emptyset$.
4. $for\ i = 1\ to\ |F'|$
   a. $for\ j = 1\ to\ f_{w_i}$
      i. $label_j = H(K_{w_i}, j||0)$ //old keyword-specific keys used.
      ii. $val_j = WMap[lable_j] \oplus H(K_{w_i}, j||1)$
      iii. $parse\ val_j\ as(\langle id_j, tf_{ij}\rangle, dlabel_j)$
      iv. $add\ id_j\ to\ set\ ResultSet.$
      v. $delete\ WMap[label_j]\ and\ re-insert$ using new keys.
      vi. $compute\ Rank\ of\ id_j$: $score_{ij} = tf_{ij} * n/f_{w_i}$, store it.
5. Compute sum of $score_{ij}$ for each of resultant file identifier in $ResultSet$
6. Sort $ResultSet$ & send encrypted files to client based on ranks.

63.4 Security and Performance Analysis of Proposed Scheme

63.4.1 Security Analysis

In our proposed scheme, we achieve the *document confidentiality* by encrypting it using a symmetric key encryption scheme like AES with a secret key K_1. The *Index* and *Query Confidentiality* is ensured as follows. The outsourced index is built employing pseudorandom functions and keyed hash functions which are based on keyword indexes rather than keywords and using random seeds during every subsequent operation like search and update in addition to initial setup operation. Hence our outsourced dictionaries like $DictW$, $DictD$, $and\,DictV$ which store label-value pairs accessible through keyed hash functions, ensure the index confidentiality. Our designed index is much secure when compared to the inverted index-based schemes or vector space model-based index like in [5, 15, 16], as the keyword-document relationship and distribution are hidden from server. And more importantly, our scheme does not allow re-run of old queries onto newly inserted documents as it invalidates the queries after search is over. Hence we achieve both forward and backward privacy which ensures both index and query confidentiality. Our scheme also achieves the *Query Unlinkability*, by using fresh random numbers every time in conjunction with keyword being searched as input to the key generation function.

63.4.2 Performance Analysis

To evaluate the performance of our proposed scheme, we implemented our semantic search algorithm using Python language with Word2Vec model being implemented in Gensim framework. We used RFC dataset for testing purpose to retrieve documents out of 8000 RFC files, based on semantic expansion of search keywords using Word2Vec model. The time it takes to set up the index involves generating label-value pairs and hashing them into a hash table using a keyed hash function, and setting up $DictDandDictV$. As there are around 5 lakh distinct keywords in RFC dataset, the index building takes significant amount of time as shown in Fig. 63.2. However, after the index is built, the search operations can be carried out in constant amount of time due to hash-based indexing structure. In schemes like [12, 13, 15], the inverted index or tree-based index are populated with precomputed relevance scores which have a heavy performance cost during update operations. Whereas in our proposed scheme, the relevance score calculations are done partially at client side and outsourced after randomization, which have a significant efficiency improvement on update operation. We also analyzed the amount of time taken to build trapdoors for user queries, as shown in Fig. 63.3. The amount of time taken to build trapdoors include time needed to compute keyword specific keys and time needed for expansion of query keywords using similarity-based expansion of Word2Vec mode. As

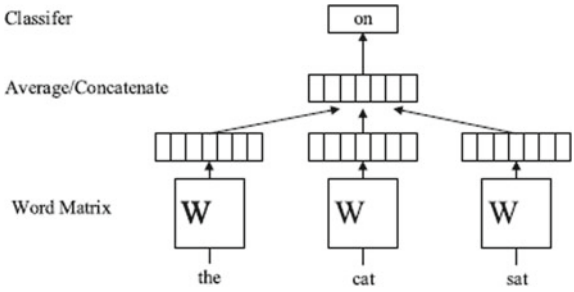


Fig. 63.1 Word2Vec framework as in [20]

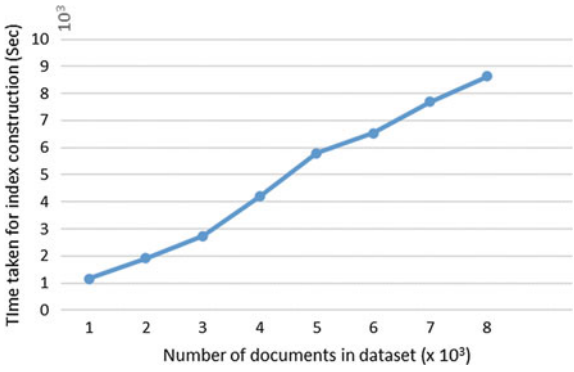


Fig. 63.2 Time taken for index setup with varying size of datasets

shown in Fig. 63.3, a major amount of time is needed for semantic expansions of query. The search operation is a sublinear cost operation as the hash-based index is searched for entries s_w times, where s_w is number of documents containing keywords semantically close to searched keyword w. The search performance of our proposed scheme is demonstrated in Fig. 63.4, with varying number of documents. The trapdoor building and search operations cost less time when compared to [13–16], as these schemes heavily use complex matrix multiplications during these two steps and also query vectors are of same size as that of document vectors. This results in heavy vectors for queries too, but our scheme utilizes trapdoors of size proportional to number of keywords in semantic expansion of query. A detailed comparison of our proposed scheme with existing semantic-aware SSE schemes is given in Table 63.1 along with asymptotic analysis of these schemes.

63.5 Conclusion

In this paper, we proposed a privacy preserving semantic-aware multi-keyword dynamic search scheme which returns documents to the user ranked as per relevance of keywords to documents, based on semantically expanded query keywords. We employ Word2Vec model to ascertain semantically closer terms to query keywords

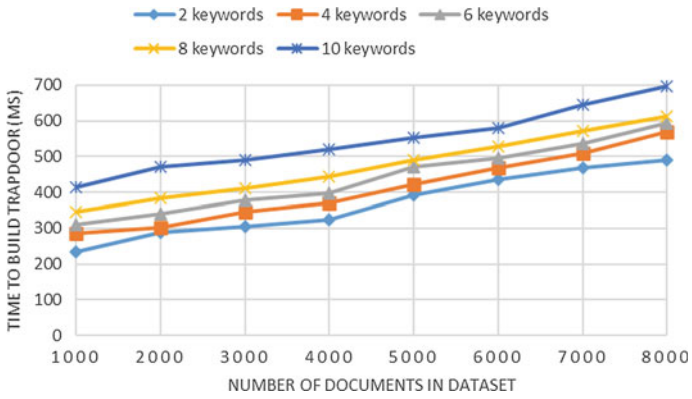


Fig. 63.3 Time to build trapdoors of varying size of queries

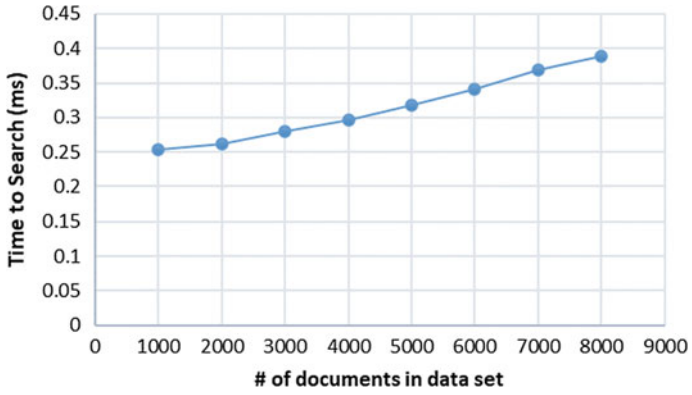


Fig. 63.4 Time to perform semantic search over different size of datasets

to match users' search intention. We use a novel hash-based secure indexing mechanism which prevents most common types of attacks such as file injection attacks. The index stores label-value pairs for each of keyword-document pairs. Upon search for a keyword, the trapdoor received by server is semantically extended to expand query and retrieves matching documents. The proposed scheme is the first scheme which reduces the size of trapdoor and efforts in search process as it does not involve either a vector-based representation of document and queries or a hierarchical indexing structure and also addresses security concerns.

Acknowledgements This work was supported under the Visvesvaraya PhD Scheme (Electronics and IT), with unique awardee number MEITY-PDH-1597. (Visvesvaraya PhD scheme: File Number PhD-MLA-4(63)/2015-16) by Ministry of Electronics and Information Technology, Government of India.

References

1. Song, et al.: Practical techniques for searches on encrypted data. In: Proceeding 2000 IEEE Symposium on Security and Privacy. S&P 2000. IEEE (2000)
2. Goh, E.-J.: Secure indexes. IACR Cryptol. EPrint Arch. **2003**, 216 (2003)
3. Cao, N., et al.: Privacy-preserving multi-keyword ranked search over encrypted cloud data. IEEE Trans. Parallel Distrib. Syst. **25**(1), 222–233 (2013)
4. Wang, B., et al.: Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud. In: IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE (2014)
5. Mikolov, T., et al.: Efficient estimation of word representations in vector space (2013). arXiv: 1301.3781
6. Etemad, M., et al.: Efficient dynamic searchable encryption with forward privacy. In: Proceedings on Privacy Enhancing Technologies **1**, 5–20 (2018)
7. Moataz, T., et al.: Semantic search over encrypted data. ICT 2013. IEEE (2013)
8. Wang, C., et al.: Achieving usable and privacy-assured similarity search over outsourced cloud data. In: 2012 Proceedings IEEE INFOCOM. IEEE (2012)
9. Fu, Z., et al.: Multi-keyword ranked search supporting synonym query over encrypted data in cloud computing. In: 2013 IEEE 32nd International Performance Computing and Communications Conference (IPCCC). IEEE (2013)
10. Sun, X., et al.: Privacy preserving keyword based semantic search over encrypted cloud data. Int. J. Sec. Appl. **8**(3), 9–20 (2014)
11. Church, K., et al.: Word association norms, mutual information, and lexicography. Comput. Linguist. **16**(1), 22–29 (1990)
12. Xia, Z., et al.: A multi-keyword ranked search over encrypted cloud data supporting semantic extension. Int. J. Multimed. Ubiquitous Eng. **11**, 107–120 (2016)
13. Fu, Z., et al.: Semantic-aware searching over encrypted data for cloud computing. IEEE Trans. Inf. Forens. Sec. **13**(9), 2359–2371 (2018)
14. Dai, H., et al.: Semantic-aware multi-keyword ranked search scheme over encrypted cloud data. J. Netw. Comput. Appl. **147**, 102442 (2019)
15. Hsieh, F.-J, et al.: Semantic multi-keyword search over encrypted cloud data with privacy preservation. In: 2019 IEEE 90th Vehicular Technology Conference. IEEE (2019)
16. Dai, X., et al.: An efficient and dynamic semantic-aware multikeyword ranked search scheme over encrypted cloud data. IEEE Access **7**, 142855–142865 (2019)
17. Zhang, et al.: All your queries are belong to us: the power of file-injection attacks on searchable encryption. In: 25th {USENIX} Security Symposium (2016)

Chapter 64
A Thorough Study on Weblog Files and Its Analysis Tools



N. Vanitha and M. Suriakala

Abstract A web server log is a data file created by a server. Many servers collect data about the information requested for the web server, examples are date, time, client IP address, recommendation, user agent, and so on. This data will help identify specific issues on the web. Log files are stored in the web server, web proxy server, and the client browser. This paper provides a detailed overview of weblog files, analytics tools, and reviews of existing research.

64.1 Introduction

The fast development of communication technology has made the dissemination of information very important. Web Log files are used to track any communication transactions about the multiple user behavior for the specific period, so that it enables as to understand the deep structure of website. The structure of the website can be understood using the Web Log files by recording the users navigation. Weblogs are effectively to clean up certain amounts of useful information. Properly compiled and illustrated log information provides statistics analysis of some important technical information regarding the server's load usage, unusual transactions, or incomplete and failed requests. It also helps with business decisions and website development.

N. Vanitha · M. Suriakala (✉)
Government Arts College, Nandanam, Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_64

64.2 Web Usage Mining

Web application mining refers to the extraction of useful knowledge from web data that includes user information. A weblog files found in three locations:

(i) web servers, (ii) web proxy servers, and (iii) client browsers.

Server Page Logs: These logs usually provide usage data, but there exist two main drawbacks are:

(a) Logs: These records contain important personal information so server owners can keep them closed.
(b) It does not record uploaded cached pages.

Proxy Page Logs: A proxy server handles the HTTP requests and sends it to a web server and then returns the results delivered by the web server. Two disadvantages:

(a) Proxy-server construction is not easy to do.
(b) Request interception is low.

Client-side logs: Test a website of an existing web browser. HTTP cookies may also be used. Two disadvantages:

(a) Need special software must install it.
(b) The operating technique is difficult to achieve compatibility with different operating systems.

64.3 Web Servers

A weblog file is a file created by a web server. Every time a visitor requests any file from the website of their requests, such as a page, image, or video, the current log file is added to the file. It contains information about the user such as time interval, URL, and IP address. There are mainly three types of log file formats used by most servers. They are common log file format, integrated log format, and multiple access logs. Four types of server logs are access log file, agent log file, error log file, and referrer log file.

Access Log: It is one of the major weblog servers; it will record each click, hit, and access of the users.

Agent Log: The file is used to record the information about the user's browser, browser version, and operating system.

Referrer Log: It is used to record the referrer log that a user came from a particular website by using the user's page link.

There is an enormous increase in web users, and everyone wants to reach the right information in a minimum period. The searches performed by the user in the browser are concentrated on the server's weblog data. Valuable information and knowledge live on this data but extracting them is not easy.

64.4 Tools

Weblogging is the process of identifying information on user browsing habits. Web mining is classified into three basic class—web content mining, web architecture mining, and web application mining. There are various tools available that can perform web application mining using web access logs as input and generate reports as output. Some of the web analytics tools are.

Though the tools are more effective, each tool has its limitations and features as follows as the table.

Tools	Primary features	Limitations
Weblog expert	Weblog is a strong log analyzer. It will provide information about the site's visitors, site routes, pages reference, search engines, user agent, and more	Recovering the HTML output can be difficult. Some log file formats are not supported
Logz.io	Logz.io is cloud-based log management. Users can create alerts for log messages and get notified via email	No dashboard with logs and metrics
Sumo logic	It allows analyzing petabyte data and learning to quickly detect patterns	Expensive, missing monitoring
Loggly	Loggly is a cloud based, user-friendly log tracking, and records analysis platform using built-in and integrated collaboration tools It allows us to track events, functions, metrics, and traffic	Rules are restrictive, though they are negotiable via custom plans
Logstash	It is the free open source for processing log data. It parses recordsfrom various sources and used in conjunction with elastic search and Kibana	Lower Performance, high resource usage
Graylog	It provides a quick warning of cyber threats. It has a simple and intuitive UI that enables us to analyze, alert, and report data	Visualization is limited
Paper trail	A paper trail is a log analyzer that automatically scans the records to provide real-time insights. It features real-time searching to quickly identify problems	Mostly output is text-based

(continued)

(continued)

Tools	Primary features	Limitations
Splunk	It is a popular log tracking and analysis site. It also collects, stores, indexes, communicates, displays analyzes, and reports on machine-generated data. Splunk allows real-time alerts	Setup costs and complexity are high
Solar winds log analyzer	It will enable us to do root cause analysis with the help of log monitoring tools and provide a real-time log stream	Limited visualization, configuration, very pricey
Log DNA	It is a cloud service used to analyze log files. It has real-time search, filtering, and debugging capabilities, a robust mechanism to help connect problems to their root cause	Limited visualization
Go access	Go access track web server records in real time using a simple command-line dashboard and analyze traffic metrics	Harder to set up, only basic tracking mechanism
Screaming frog log file analyzer	Allows checking search engine bots, detect crawled URLs, and analyzing the data, etc.	Limited functionality, crawl a limited number of URLs for the free version
Apache log viewer	Apache logs viewer can easily filter and analyze Apache/IIS/Engine log files. The search and filter function is configured. Custom record formats are also supported	Though it is free for unlimited use, some features are locked

Most of the tools are traffic analyzers. Their features are limited to generating reports on website traffic: The user visits and page views, user statistics, operating systems, and web browser and page not found. Some tools analyze user statistics and specific statistics about individual users. There are many business analytics tools available, but most of them have speed limits, which are expensive, inflexible, and difficult to maintain.

64.5 Literature Review

Much research has so far focused to solve this issue on analyzing weblog data. The methods discussed by different researchers used various measures and techniques.

Zhang et al. proposed that the use of association rules analysis in weblog mining is of great importance for achieving a personalized website and personalized service. In

this study, the matrix a priori algorithm is used for weblog mining for sogou weblog mining and footwear websites compared to the four association rules algorithms [1].

Koutsoukos et al. show that a new method is proposed for session identification in weblog data based on ambiguous the fuzzy c-means clustering algorithm. The novelty of the approach lies in launching obscure partition matrices using subtractive c1ustering technology [2].

Cao et al. this paper described an unknown detection system on machine learning for weblog files. The system will be used for analysis log files. This paper first preprocesses log files by using decision tree algorithms based on patterns based on the ruleset, then modeling the normalized data using HMM to create a model that is used to specify the default data state as an abnormal detector [3].

Anuradhab et al. this study is based on a review of 14 college's weblog files. After the estimation, we define the cluster more accurately and present the most accurate results and prediction model based on the Markov chain and Bayesian theorem [4].

Zhu et al. this study proposes an advanced method of F&D to expedite the mining process of WFPS from weblogs, in which weights are allocated by the residence time to find not only frequent visits but also important page collections [5].

Latib et al. the results presented here will facilitate improvements an exploration of weblog files for analysis of weblog files using Hadoop [6].

Agichtein et al. this study explored the use of combining the noisy implicit feedback obtained in the actual web search system to improve search rankings [7].

Kumar et al. this paper is to track visitors' search behavior by using the tool and maintain the website according to this behavior. In model discovery, the association will use different mining techniques such as rule, classification, clustering, and continuous sampling technique to find important information [8].

Patel et al. this research work includes the pre-implementation phase and web application mining, which can be used in industry and application-oriented organizations. Used customized weblogging pre-processing rather than the traditional approach that reduces the size of the original weblog file. The development will show up in time and accuracy [9].

64.6 Conclusion

Weblog files play a significant role in web mining research because it contains user information and user behavior can be extracted from it. As web traffic has already dominated the Internet, massive weblogs are being generated ceaselessly. It is essential and meaningful for operators to mine valuable information and knowledge from the log data. However, the state-less feature of HTTP and increasing dynamics and complexities of web services bring a challenge to web mining in weblogs. Techniques need to be improved to reduce delays in accessing web pages. Weblog mining, however, still has many problems to be solved and in-depth studied. In conclusion, it is hoped that this paper will provide some depth into the tools and review of

existing works on log analysis and explained how and why these weblog analyses are important.

References

1. Zhang, H., Song, W., Liu, L., Wang, H.: The application of matrix apriori algorithm in web log mining. In: IEEE 2nd International Conference on Big Data Analysis (2017)
2. Koutsoukos, D., Alexandridist, G., Siolas, G., Stafylopatis, A.: A new approach to session identification by applying fuzzy c-means clustering on web logs. In: IEEE Symposium Series on Computational Intelligence (SSCI) (2016)
3. Cao, Q., Qiao, Y., Lyu, Z.: Machine learning to detect anomalies in web log analysis. In: IEEE International Conference on Computer and Communications (2017)
4. Anuradhab, H.K., Solankic, A.K., Krishna, K.: Progressive machine learning approach with webastro for web usage mining. In: International Conference on Computational Intelligence and Data Science (ICCIDS 2019)
5. Zhu, Z., Xu, Q.: Fast mining frequent page sets from web log by filtering strategy. In: Proceedings of the 2nd International Conference on Computer Science and Application Engineering, October 2018
6. Latib, M.A., Ismail, S.A., Yusop, O.M., Magalingam, P., Azmi, A.: Analysing Log Files For Web Intrusion Investigation Using Hadoop ICSIE '18
7. Agichtein, E., Brill, E., Dumais, S.: Improving web search ranking by incorporating user behavior information, SIGIR (2018)
8. Kumar, M., Meenu: Analysis of visitor's behavior from web log using web log expert tool. In: International Conference on Electronics, Communication and Aerospace Technology ICECA (2017)
9. Patel, R., Panchal, K., Rathod, D.: Efficient log mining from web server using clustering technique. J. Emerg. Technol. Innov. Res. (2016)

Chapter 65
Analysis of Heart Rate Variability Using Different Techniques of Breathing



R. Kalpana, C. R. Sai Krishna, Nihar Jajodia, and M. N. Mamatha

Abstract Heart rate variability plays a major role in the signal processing. The objective of the study is done by considering two important parts of the nervous system of the brain is known as a sympathetic and parasympathetic which are in turn useful for analysis of heart rate variability. These are analyzed using the lower frequency and higher frequency bands. The powers of this ratio will in turn helps to know the quantification of sympathovagal balance. Use of nonlinear methods such as Poincare plot, approximate entropy and correlation dimension has demonstrated in correctly enhancing information from biomedical signals pranayama or breathing exercise leads to better sympathovagal balance from Poincare plot analysis. Sympathetic and parasympathetic nervous system.

65.1 Introduction

Pranayama or yogic breathing has long been reported to spice up health and reduce the implications of stress and strain on the body and modern-day lifestyle has made people to induce mental pressure, especially within the metropolitan cities, due to increasing expectation and thus leading to depression it has been known that health beneficial effects of Pranayama are carried over through by meditation and mainly through improvement of autonomic functions. The bulk of individuals are littered with stress and strain. This could be significantly reduced by practicing certain yoga techniques for our improvement of health benefits, especially for the vital

R. Kalpana (✉) · C. R. Sai Krishna · N. Jajodia
Department of Medical Electronics Engineering, B.M.S College of Engineering, Bangalore 560019, Karnataka, India
e-mail: rk.ml@bmsce.ac.in

M. N. Mamatha
Department of Electronics and Instrumentation Engineering, B.M.S College of Engineering, Bangalore 560019, Karnataka, India
e-mail: mamathamnbms.intn@bmsce.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_65

signals of our body. This successively enhances our condition of health. These chronic stresses, affects health and it produces various forms of bodily dysfunction. Therefore, the incidence of stress-induced disorders [1]. That has increased to an oversized within the past few years that particularly increases the cardiovascular abnormalities. It has been studied that several stress and strain-induced disorders occur thanks to autonomic imbalance, the imbalance between sympathetic and parasympathetic functions [1] brain and heart for people that are full of depression and other mental, and chronic diseases. Due to these lifestyle this has a bearing on people that are laid low with diabetics, hypertension obesity, and other stress related syndromes which successively affects the youngsters and young adults at greater measure. However, very less study has been made till date to assess the issues of autonomic balance achieved by practice of pranayama [2] or yogic breathing on reduction of prevalence of such metabolic disorders [1]. There are descriptions from our practice that tells us on how slow pranayama breathing is more practical in improving autonomic functions than the first pranayama breathing. With the use of kapalpathi yoga practice we can reach balance of the sympathetic activity. Though there was an earlier study that reported some extent of dominance in autonomic functions without significant difference in cardio-respiratory functions within the groups practicing left nostril breathing and right nostril breathing exercises. Cardiovascular functions is healthier ensured by strong sympathovagal balance as system is primarily under the control of autonomic system Therefore, effective sympathovagal homeostasis is that the foremost objective of various physical exercises and yogic therapies which are aiming at providing healthy psychological and physical life. Recently spectral analysis of pulse rate variability has been documented as a sensitive measure of autonomic function and dysfunction in health and clinical disorders. Though the results of right and left nostril breathing on sympathetic activity and blood Pressure are known, the implications on variability of heart rate and cardiovascular abnormalities risks don't seem to be yet investigated. Therefore, during this study we've assessed the difference within the results of right nostril breathing and left nostril breathing practice on sympathovagal balance, heart rate variation and cardiovascular abnormalities assessed by spectral analysis of variation of heart rate in young adults with the use of nonlinear analysis (Fig. 65.1).

65.1.1 Data Acquisition

Data was collected from 10 people ageing between 18 and 24 years. This was done in Swami Vivekananda center for Yoga Anusandhana Samsthana or S-VYASA research center. Where the ECG of 10 Male subjects were recorder for four different breathing techniques or 'pranayama'. Data was recorded using an electrocardiogram (ECG).The data acquisition process is done four different days for the four different exercises. The ECG is then segmented into different parts as per the activity in which the subject was involved. These parts contain initial baseline (5 min), three neurocognitive tests before the breathing Exercise (15 mins), three repetitions of the breathing

exercise (ANB, BA, LNB, and RNB), and three post breathing neurocognitive tests (Fig. 65.2).

65.1.2 Data Conversion

Data conversion is required to convert ECG waveforms to text and numeric tables for analysis. This is done using a software named Acknowledge software. Acknowledge is a special software that has many user-friendly features in it. User could select any of them for his analysis, few of them are averaging, sample rates, automation and duration's in acquisition modes. Graphical windows are completely customized by allowing multiple views of the identical data. Transformations can be done by rearranging tool bars and by adding custom buttons. There may well be change the color or size of grids, waveforms, backgrounds, text, and more. Acknowledge can output a variety of stimulus paradigms for constant voltage stimulation and constant current. It accepts digital events from computer game environments and stimulus presentation systems like Presentation, E-Prime, and Super Lab and categorizes them for further future analysis and they'll also easily configure arbitrary waveforms, single pulses, and pulse chains for stimulation.

65.1.3 Data Processing

ECG processing involves extraction of parameters like pulse Variability, mean RR interval, LFnu, HFnu, LF/HF, etc. this can be done employing a software called kubios HRV Using the software ECG analysis can be done effectively for finding the suitable parameters used for the analysis such as detection of heart rate, using the amplitudes of R peaks, the software has an ability to accept other file-formats, filtering and trend analysis and automatic detection of QRS components The computation and analysis of Time as well as frequency domain of heart rate variability analysis and also many nonlinear-parameters can be obtained using the proposed software. There are different flexible analysis settings are adopted for optimization of varied datas. In addition to that the respiratory frequency can also computed for our analysis (Fig. 65.3).

65.1.4 Data Analysis

SPSS software was used for Analyzing the Results Obtained from the kubios. There was a comparison done on different breathing exercise through nonlinear methods. Data was processed using SPSS Version 14.0. The parameters suitable for our study is analyzed using analysis of variance and T test is also applied and the value of p

< 0.05 was taken for calculation is significant. The related hemodynamic variables such as ratio between low to high frequency and Standard deviations ratio can also calculated

65.2 Linear and Nonlinear Methods

65.2.1 Linear Methods

HRV could be analyzed in time or frequency domain using linear algorithms. This technique assumes. ECG is a stationary signal. Heart rate Variability [1] techniques using linear algorithm [3] are computationally simple. Information content in the RR interval cannot be identified or analyzed through Linear HRV analysis as cardiac functions are complex and nonlinear owing to multiple control loops. Hence there are limitations in the linear system. For the calculation of heart rate variability Time domain features are the most important concern as they calculated from statistical methods and R-R intervals and they can also be correlated with one another. Frequency domain indexes are more elaborated indexes supported spectral analysis, mostly accustomed evaluate the Contribution on HRV of autonomic system (VLF, LF, HF, HF/LF ratio) (Fig. 65.4).

65.2.2 Non-linear Methods

Nonlinear analysis are newly introduced methods to measure rate of variability of heart. They include nonlinear parameters such as, Correlation dimension, Poincare geometry, Hurst exponent, Lypnov exponent, and Approximate Entropy [4, 5]. All complex interactions of electrophysiological, hemodynamic, and humoral variables additionally as by the autonomic and central nervous regulations are understood. These techniques are shown to be powerful tools for characterization of complex biological systems. RR interval contains information about different physiological states thanks to different interacting systems. Hence we get to understand about more complicated physiological data that provides knowledge about vital sign, mental pressure, and respiration. The complex interplay between Parasympathetic and sympathetic [6] stimulation are often detected by Nonlinear pulse Analysis (Fig. 65.5). These nonlinear analysis and methods are applied in Bio-signal Processing and Applications [6–9].

65.3 Results

The following differences in the means of the nine parameters considered in this study to evaluate sympathovagal balance are tabulated below. The nine parameters are (Fig. 65.6): The datas before Pranayama and After Pranayama are tabulted in the following with the subjects taken for Analysis Tables (65.1 and 65.2).

1. Mean RR interval
2. Mean Heart Rate
3. LFnu
4. HFnu
5. LF/HF
6. SD1
7. SD2
8. SD2/SD1
9. Approximate Entropy

Table 65.1 Data before pranayama

Parameters	Sub1	Sub2	Sub3	Sub4	Sub5
Mean HR	71.07	75.59	66.04	83.35	70.47
LF	64.5	48.8	63.9	89.3	65.4
HF	35.5	51	36	10.7	34.4
LF/HF	1.81	0.95	1.74	8.37	1.9
SD1	26	15.6	31.7	59.8	28.1
SD2	71.2	56	107	184	61
SD2/SD1	2.73	3.5	3.37	3.07	2.17

Table 65.2 Data after pranayama

Parameters	Sub1	Sub2	Sub3	Sub4	Sub5
Mean HR	74.55	64.27	68.04	85.35	70.96
LF	45.1	44.8	41.8	60	44.8
HF	54.8	55	58.2	40	54.8
LF/HF	0.823	0.814	0.718	1.499	0.817
SD1	25	31.4	23.2	35.4	30.3
SD2	76.4	60.7	46	69.2	42
SD2/SD1	3.056	1.93	1.98	1.95	1.38

65.4 T Test

T-test is a useful hypothesis applied between the means of two groups which might be related in certain features. To conduct a test with three or more variables, one must use an analysis of variance. To conduct a test with three or more variables, one must use an analysis of variance. T-test was done for various breathing exercise in order to determine the significant nonlinear components. The mean heart rate, SD1, SD2, SD2/SD1, and Approximate Entropy were found to be significant. The mathematical equivalence of LF/HF in the Linear method Correlated with the SD2/SD1 [5]. Thus there was much similarity of LF to SD2 and HF to SD1. The mean heart rate, LF, HF, LFnu, HFnu, SD1, SD2, SD2/SD1, and approximate entropy was calculated for all different exercises. Only few exercises like Alternate breathing and Post Alternate breathing had significance changes with respect to most of the linear and nonlinear parameters. The ratio of standard deviation was seen reduced 1 digit lower from starting data recorded (Figs. 65.7 and 65.8).

65.5 Conclusions

Yogic breathing or pranayama practices were known to influence the neurocognitive abilities, pulmonary and autonomic functions as well as the metabolic and biochemical activities in the body. With the SD1 decreasing, the SD2 increases, and the ratio of SD2/SD1 decreases instantly after the sympathetic stimulation with concurrent change of shape in the Poincare plot. Parasympathetic activity is determined by Standard deviation1 and Sympathetic activity by Standard deviation2. Like the LF/HF the ratio of SD2/SD1 could also determine sympathovagal balance. Sympathovagal changes can be easily detected by the changes in the shape of Poincare plot mainly during the tilt position. We found that the ratio of SD2/SD1 was greater for persons before doing the breathing exercise. After the breathing was done, Ratio gradually decreased and thus maintaining the Sympathovagal balance. The proposed work aims for finding out some more nonlinear parameters such as approximate entropy, Correlation Dimension, Sample entropy, which can also give information about parasympathetic and sympathetic parameters. The above research can enhance the work for people who are suffering from heart-related as well as stress-related for leading them to live a healthy lifestyle and mental peace. The Changing Lifestyle of Man or Woman has brought more stress in them and as a result they undergo through Depression. Yogic breathing not only brings the relax state it is also a powerful Preventive measure for mental peace and Physical Wellness. Overall, we found the practice of yogic breathing safe, when practiced under guidance of a trained teacher.



Fig. 65.1 Block diagram of the Proposed work

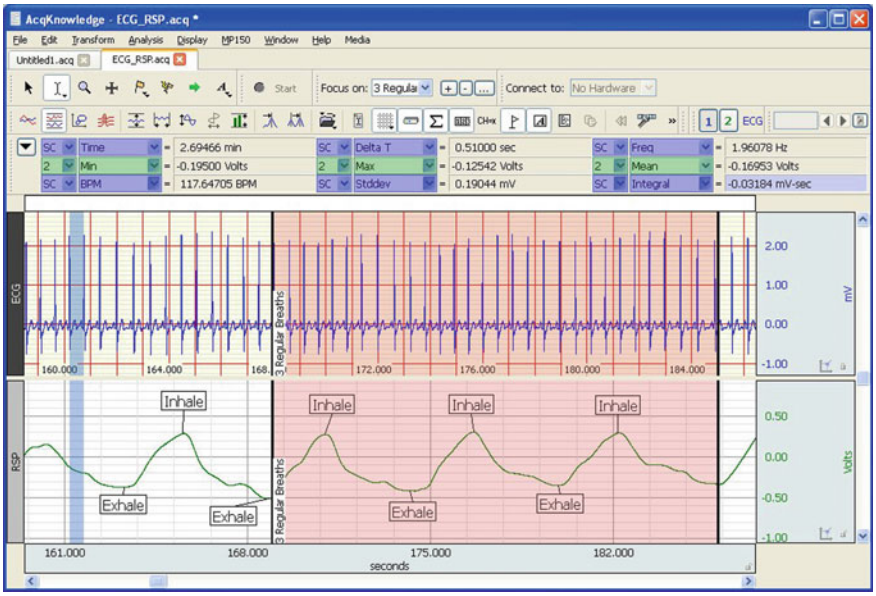


Fig. 65.2 Graphical output of the Proposed work

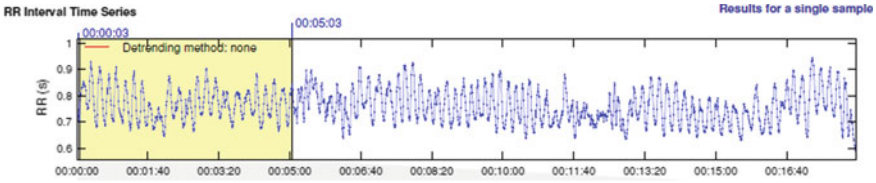
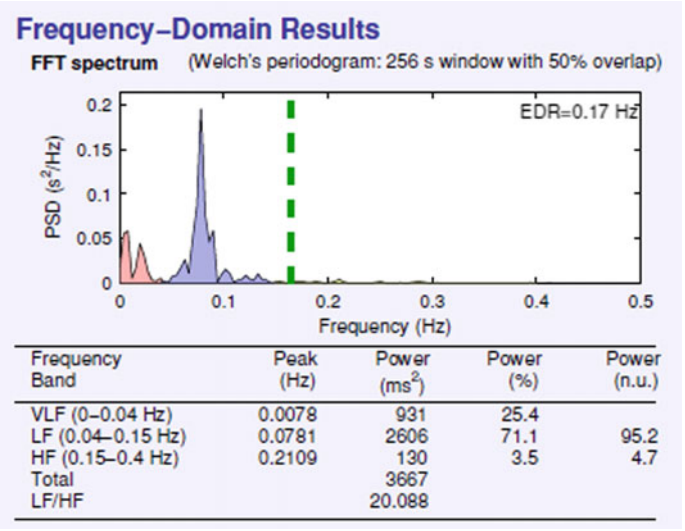
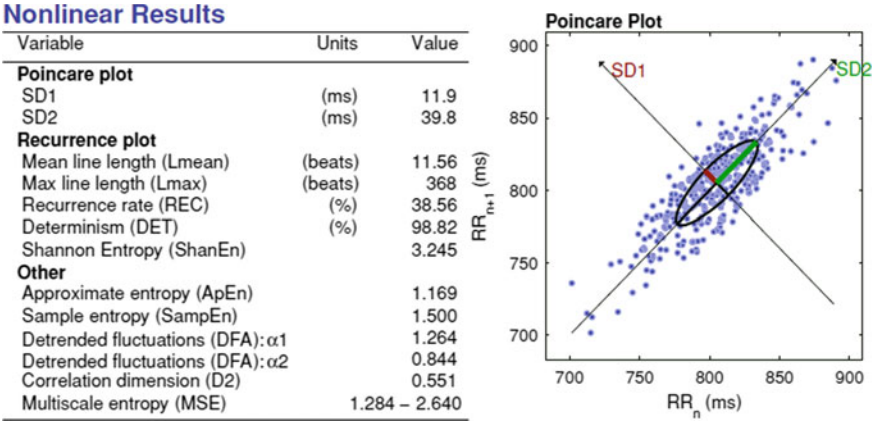


Fig. 65.3 Time Series Analysis



Frequency Band	Peak (Hz)	Power (ms^2)	Power (%)	Power (n.u.)
VLF (0–0.04 Hz)	0.0078	931	25.4	
LF (0.04–0.15 Hz)	0.0781	2606	71.1	95.2
HF (0.15–0.4 Hz)	0.2109	130	3.5	4.7
Total		3667		
LF/HF		20.088		

Fig. 65.4 Frequency Domain Analysis



Variable	Units	Value
Poincare plot		
SD1	(ms)	11.9
SD2	(ms)	39.8
Recurrence plot		
Mean line length (Lmean)	(beats)	11.56
Max line length (Lmax)	(beats)	368
Recurrence rate (REC)	(%)	38.56
Determinism (DET)	(%)	98.82
Shannon Entropy (ShanEn)		3.245
Other		
Approximate entropy (ApEn)		1.169
Sample entropy (SampEn)		1.500
Detrended fluctuations (DFA): $\alpha 1$		1.264
Detrended fluctuations (DFA): $\alpha 2$		0.844
Correlation dimension (D2)		0.551
Multiscale entropy (MSE)		1.284 – 2.640

Fig. 65.5 NonLinear Analysis

B	C	D	E	F	G	H	I	J	K	L	M
		PERSON 1	PERSON 2	PERSON 3	PERSON 4	PERSON 5	PERSON 6	PERSON 7	PERSON 8	PERSON 9	PERSON 10
BASELINE	MEAN RR	847.9	795.9	915.8	741.7	854.1	923.3	736.7	923.9	802.9	1054.1
	MEAN HR	71.07	75.59	66.04	83.35	70.47	65.37	81.7	66.08	75	57.26
	Lfnu	64.5	48.8	63.9	89.3	65.4	53	70.8	77.5	34.6	45
	Hfnu	35.5	51	36	10.7	34.4	47	29.2	22.5	65.2	55
	LF/HF	1.818	0.957	1.744	8.375	1.902	1.126	2.429	3.449	0.531	0.819
	SD1	26	15.6	31.7	59.8	28.1	36.5	17.5	71.7	28.1	51.9
	SD2	71.2	56	107.9	184.7	61.3	93.5	56.2	156.6	61.2	97.7
	SD2/SD1										
	SAMPLE ENTROPY(SampEn)	1.572	1.483	1.324	0.422	1.544	1.524	1.654	1.144	1.855	1.42
	APPROXIMATE ENTROPY(ApEn)	1.142	1.188	1.06	0.625	1.136	1.106	1.287	0.93	1.236	1.022
ANB1	MEAN RR	795.6	681.8	851.7	742.1	808.8	806.7	701.8	754	739.1	919.6
	MEAN HR	76.09	88.44	71.22	83.54	75.1	75.13	86.88	80.44	69.9	66.28
	Lfnu	94.5	96	96.6	89.5	89.9	92.5	97	95	46.7	97.5
	Hfnu	5.5	4	2.1	10.5	10	7.5	3	5	53.3	2.5
	LF/HF	17.165	23.989	28.901	8.51	8.954	12.396	32.738	19.12	0.875	39.521
	SD1	103.9	13.5	22.2	65.7	38.4	35.7	27.2	25.6	31.5	32.9
	SD2	26.1	67.5	124.5	189.6	119.8	110.6	124.9	110.4	93.8	156.6
	SD2/SD1										
	SAMPLE ENTROPY(SampEn)	0.99	0.863	0.861	0.339	0.901	1.118	0.593	0.864	1.149	0.919
	APPROXIMATE ENTROPY(ApEn)	0.928	0.854	0.823	0.55	0.839	0.989	0.682	0.835	0.98	0.835
	MEAN RR	799.5	706.8	862.9	743.6	806.7	825.8	688.3	772.8	720.1	891.5
	MEAN HR	75.64	85.35	70.42	83.2	75.14	73.37	88.3	80.1	83.76	68.48
	Lfnu	94.2	95.8	92.1	90	88.2	92.9	90.8	96.1	26.3	96.9
	Hfnu	5.8	4.2	7.9	10	11.8	7.1	9.2	3.9	73.7	3.1

Activate
Go to Settin

Fig. 65.6 Nonlinear Parameters

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BS- MEAN RR	859.630	10	97.5394	30.8447
	ANB1 MEAN RR	780.120	10	71.4596	22.5975
Pair 2	BS- MEAN HR	71.1930	10	7.98113	2.52386
	ANB1 MEAN HR	77.3020	10	7.37001	2.33060
Pair 3	BS- Lfnu	61.280	10	16.2420	5.1362
	ANB1 Lfnu	89.520	10	15.3085	4.8410
Pair 4	BS- Hfnu	38.650	10	16.1960	5.1216
	ANB1 Hfnu	10.340	10	15.3809	4.8639
Pair 5	BS- LF/HF	2.314133955	10	2.286377569	.7230160709
	ANB1 LF/HF	20.82384670	10	14.47099638	4.576130857
Pair 6	BS- SD1	36.690	10	18.5321	5.8604
	ANB1 SD1	39.660	10	26.4082	8.3510
Pair 7	BS- SD2	94.630	10	44.5798	14.0974
	ANB1 SD2	112.380	10	44.8303	14.1766
Pair 8	BS- SD2/SD1	2.701968909	10	.5959161523	.1884452336
	ANB1 SD2/SD1	3.660505478	10	1.541533720	.4874757646
Pair 9	BS ApEn	1.39420	10	.390439	.123468
	ANB1 ApEn	.85970	10	.239189	.075638

Fig. 65.7 Post right nostril breathing

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BS- MEAN RR	859.630	10	97.5394	30.8447
	Post ANB RR	818.880	10	80.4275	25.4334
Pair 2	BS- MEAN HR	71.1930	10	7.98113	2.52386
	Post ANB HR	74.3370	10	7.46173	2.35961
Pair 3	BS- Lfnu	61.280	10	16.2420	5.1362
	Post ANB LFNU	60.630	10	14.7714	4.6711
Pair 4	BS- Hfnu	38.650	10	16.1960	5.1216
	Post ANB HFNU	39.630	10	14.6075	4.6193
Pair 5	BS- LF/HF	2.314133955	11	2.169048213	.6539926430
	Post ANB LF/HF	1.909300	11	1.0698083	.3225593
Pair 6	BS- SD1	36.690	10	18.5321	5.8604
	Post ANB SD1	31.960	10	12.1728	3.8494
Pair 7	BS- SD2	94.630	10	44.5798	14.0974
	Post ANB SD2	72.150	10	28.6391	9.0565
Pair 8	BS- SD2/SD1	2.701968909	10	.5959161523	.1884452336
	Post ANB SD2/SD1	2.414306247	10	1.084664016	.3430008788
Pair 9	BS ApEn	1.39420	10	.390439	.123468
	Post ANB ApEN	1.12850	10	.091234	.028851

Fig. 65.8 Post alternate nostril breathing

References

1. Aparvinash Saoji, B.R., Raghavendra, N.K., Manjunath: Effects of yogic breath regulation: a narrative review of scientific evidence
2. Schubert, C., Lambertz, M., Nelesen, R.A., Bardwell, W., Choi, J.B., Dimsdale, J.E.: Effects of stress on heart rate complexity—a comparison between short-term and chronic stress. Biol. Psychol. **80**, 325–332 (2009). https://doi.org/10.1016/.j.biopsycho.2008.11.005
3. Shinba, T., Kariya, N., Matsui, Y., Ozawa, N., Matsuda, Y., Yamamoto, K.: Decrease in heart rate variability response to task is related to anxiety and depressiveness in normal subjects. Psychiatry Clin. Neurosci. **62**, 603–609 (2008). https://doi.org/10.1111/j.1440-819.2008.01855.x
4. Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C.M., Suri, J.S.: Heart rate variability: a review. Med. Biol. Eng. Comput. **44**, 1031–1051 (2006)
5. Dabire, H., Mestivier, D., Jarnet, J., Safar, M.E., Chau, N.P.: Quantification of sympathetic and parasympathetic tones by nonlinear indexes in normotensive rats. Am. J. Physiol. Heart Circul. Physiol. **275**, H1290–H1297 (1998)
6. Pal, G.K., Agarwal, A., Nanda, N.: Slow yogic breathing through LNB and RNB influences sympathovagal balance, heart rate variability and cardiovascular risks in adults. Department of Physiology, Jawaharlal Institute of Postgraduate Medical Education and Research, Pondicherry, India. PMCID: PMC3978938
7. Joshi, M., Bairagi, S.: Hrv analysis using linear and nonlinear techniques. In diabetic: research article. BMR Med. **2**(1), 1–6
8. Hsu, C.-H., Li, C.-Y.: Poincare plot indexes of HRV detect dynamic autonomic modulation during general anesthesia induction. Department of Anesthesiology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan
9. Mirescu, S.-C., Harden, S.W.: Nonlinear dynamics methods for assessing HRV in patients with recent myocardial infarction. Burnett School of biomedical Sciences, College of medicine, University of Central Florida, USA

Chapter 66
Minimum Time Search Methods for Unmanned Aerial Vehicles



R. Darsini, Nikita Ganvkar, Karthik Gurunathan, and Ranjita Kumari Dash

Abstract This paper analyzes and compares methods used to evaluate and obtain the best trajectories of multiple Unmanned Aerial Vehicles (UAV) that have to find a lost target as soon as possible. The analysis shows how the high-complexity of this problem, named Minimum Time Search (MTS), makes it a good candidate to solve it with different bio-inspired and probabilistic-based approaches and also elaborates some of the advantages and drawbacks of these approaches. The paper also compares and relates two different mathematical formulations that can be used to evaluate the UAV trajectories based on the mean-time to detect the targets with the sensors on-board the UAVs.

66.1 Introduction

UAVs are aircrafts that do not have a human pilot on board, and they may or may not be remotely controlled. More specifically, UAVs are essential components of Unmanned Aircraft Systems (UAS). The other components of UAS are Ground Control Stations (GCS) and its communication system with the UAVs. The UAV itself comprises of: (i) a frame (the skeleton of the UAV); (ii) a propeller (the physical actuators that hold the UAV in flight); (iii) the engines (makes the propellers rotate); and (iv) the flight controller (generates the actions that ensures the UAV's stability and correct displacement during the flight).

UAVs were originally used for military purposes (e.g. tactical reconnaissance or target tracking), but their applications have been extended to other fields (e.g. inspection, monitoring, surveying, mapping, aerial imaging and video recording). The paper mainly focuses on Search And Rescue (SAR) operations, where the UAVs have to determine the location of one or several lost targets. The main advantage of UAVs over piloted vehicles in these situations is the elimination of the human risk

R. Darsini (✉) · N. Ganvkar
Department of Computer Science and Engineering, PES University, Bengaluru, India

K. Gurunathan · R. K. Dash
Department of Computer Science and Engineering, PES Institute of Technology, Bengaluru, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_66

factor in the mission. Moreover, UAVs can offer more endurance and range than their manned counterparts, as the Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis of UAS-based applications presented in [5] shows.

In terms of SAR operations, time becomes an essential factor in the mission. This is especially true for missions which have a dire aftermath if the target is not found on time. For such scenarios, the target has to be found in the least possible time by a team of UAVs in a SAR mission.

MTS approaches determine UAV's trajectories that cut down the target detection time by considering the available information about the SAR scenario. MTS is often optimized using bio-inspired algorithms and probabilistic-based optimization approaches. A prime bio-inspired algorithm explored in the paper is Ant Colony Optimization (ACO), which can also incorporate a problem-specific heuristic to increment the chances on generating overall good solutions quickly. Other bio-inspired approaches are Evolutionary Algorithms (EA), while probabilistic-based optimization techniques are the Bayesian Optimization Algorithm (BOA) and Cross Entropy Optimization (CEO). Using such approaches, MTS can be used as a framework for single and multi-UAVs SAR mission.

This paper analyses bio-inspired and probabilistic-based optimization algorithms well suited for multi-UAV MTS, highlighting their main advantages and limitations. Besides, after the literature review presented in Sect. 66.2, the optimization approaches are described in Sect. 66.3, the mathematical models for evaluating the solutions are analyzed in Sect. 66.4. The conclusion and future work are presented in Sect. 66.5.

66.2 Literature Overview

This section describes the main bio-inspired and probability-based optimization approaches for MTS. The main advantages and limitations of these approaches are summarized in Table 66.1.

A prime example of the implementation of bio-inspired approaches is [9], where the underlying goal is to locate a lost target as quickly as possible using a fleet of UAVs. This is done by finding the path required to fly quickly to the regions where there are high chances of finding the target (using ACO). The approach in [9] also includes a new heuristic for MTS where the target's spatial and probabilistic properties are exploited while building the UAV trajectory. The performance of this algorithm is tested against various scenarios and is compared with other MTS based algorithms (ad-hoc heuristics in [6], CEO, BOA and Genetic Algorithms- GA). The comparison clearly shows that the approach in [9] provides better solutions than the other algorithms used for MTS.

The feasibility of ACO for MTS is further explored in [11]. This is done by complementing the original MTS heuristic presented in [9] with new terms that account for important constraints such as maintaining the UAV's communication with the GCS and collision avoidance among the UAVs. This algorithm is tested

Table 66.1 Comparison of closely related works based on MTS

Author	Advantages	Limitations
Perez Carabaza et al. [9] MTS-ACO	High level straight segmented trajectories are obtained by exploiting spatial properties of the problem itself	Restricts the UAV mainly towards regions with closer and higher probability of target presence
Perez Carabaza et al. [11] MTS-ACO	Computing UAV trajectories after including the possibility of collision and loss of communication loss with GCS	Synthetic scenarios do not mirror real world situations
Perez Carabaza et al. [7] MTS-ACO	It shortens the detection time of a given target by using a real-coded ACO_R that compares the UAV performance using a complex dynamic model	Artificial ants of ACO_R follow heuristic based methodology similar to patrollers and do not use inputs from the pheromone table
Perez Carabaza et al. [8] MTS-EA	It includes a complex dynamic model for the UAV and evaluates UAV performance with multiple criteria	Computational time increases due to the calculations of four different optimization criteria
Perez Carabaza et al. [10] MTS-EA	It combines intelligence and geographical information to construct target probability models that will be used in real-world missions. It simultaneously optimizes the UAV trajectory and camera movement	Model does not handle probabilistic behavior for multiple targets
Bourgault et al. [1] MTS-Greedy MTS-Control Optimization	It introduces the continuous mathematical formulation of the single UAV MTS problem with arguments that make it applicable to simplified real-world problems	Being a grid based approach (as the others), it is difficult and computationally costly to increase the search area
Lanillos et al. [3] MTS-CEO	It focuses on the discretization of the evaluation of the MTS, applicable to only one UAV at a time instead of including multiple constraints and increasing the uncertainty	Targets with different probability location distributions are out of scope
Lanillos et al. [4] MTS-BOA	Expressions of single-agent MTS problem are extended to accommodate the multi-agent variant	BOA computational cost makes it a slow approach for trajectories with many segments
Meghjani et al. [6] MTS-heuristics	UAVs are driven to overfly highest probability areas by ad-hoc heuristics specially designed for MTS	The heuristics have only been tested with static target scenarios

with pre-defined scenarios and is successfully compared against extended CEO and GA-based MTS algorithms. However, neither the scenarios in [9] nor [11] completely mirror real-world situations as they are theoretically built, to show the superiority of ACO over other algorithms. Besides, both approaches generate trajectories defined by straight-line segments, which can only be followed by rotatory-wing UAVs.

In order to generate curved trajectories that can also be followed by fixed-wing UAVs, the approach in [7] makes use of a complex dynamical model of the UAV, whose reference signals are optimized with ACO_R (an ACO algorithm specifically designed for real-coded domains). Besides, in order to use a novel MTS heuristic in ACO_R, the approach in [7] includes a certain percentage of ants that build their solutions using the proposed heuristic.To deal better with real-world scenarios, the approach in [8] uses an EA for multi-objective constrained problems which includes several criteria: Non-Flying Zones (NFZ) and collision avoidance, fuel consumption, expected detection time, myopia reduction index, and trajectory smoothness. Furthermore, [10] replaces artificial scenarios with real-world ones and it allows to simultaneously optimize both the UAV trajectory and the camera pose. This work also presents three detailed real-world inspired scenarios and includes the impact of different variables (e.g. search area, grid size, and UAV flying height).

To model the uncertainty about the target location in each scenario, all the approaches in Table 66.1 explore the connection between MTS and Bayes theorem. The introductory work in [1] focuses on the condition that a single UAV searches for a non-evading single target, which may be static or in motion. To be able to use Bayes theorem, unknown quantities are cast as random variables, whose information are expressed as probability density functions (PDF). In the MTS problem, an unknown quantity is the target location. Bayes theorem is used to combine new probability information with prior probability information to form a new PDF. In the MTS problem, the Bayes theorem combines the detection measurements provided by the sensors onboard the UAVs with the probability of locating the target just before the observation is obtained to form a new updated PDF. Besides, if the target is in motion, the state of the PDF is also predicted by increasing discrete time intervals of the UAV time of flight. The predicted and updated PDF is used to calculate the *expected time* of target detection, which is optimized through different approaches to obtain the best references for the UAV and their associated trajectories to find the target.

To give a fresh outlook in solving the MTS problem, CEO is applied in [3] to compute the best N number of actions for several UAVs that has to find a single non-evading mobile target in the minimum possible time. This work also proposes to optimize the discounted-time probability of target detection for MTS.

A more extensive use of Bayesian Theory is put to test in [4], which in fact, utilizes the Bayes theorem twice: (i) for evaluating the expected time of target detection in order to compare the routes of the agents and (ii) within an optimization algorithm that exploits Bayesian theorem and probability independence to solve the problem. Besides, it formalizes the MTS problems as a Partially Observable Markov Decision Process (POMDP) using the target initial location belief, and probabilistic models for the target motions and sensors. The approach in [4] also extends the evaluation

of the expected time of target detection for multi-UAV MTS problems. This shows that minimising the time taken to find the target increases the probability of finding the target itself, which is successfully tested over five scenarios.

An alternative approach is presented in [6], which introduces several constructive heuristics designed ad-hoc for the MTS problem of a single static target. These heuristics can quickly calculate the UAVs trajectories, avoiding the optimization process of MTS-related criteria that occurs in the other approaches.

Finally, it is worth noting that the previous analyses illustrates how bio-inspired and probability-based approaches can be used to solve different MTS problems successfully. The promising result of ACO, a member of swarm optimization, makes it interesting to analyze the behavior of these types of approaches for MTS. An example is the work in [2], a behaviour-based approach for searching lost targets, with a reduced need of human control. This is done by providing personalization parameters to the UAV, on top of the existing hive mind algorithm (where the UAVs share behaviours and are controlled by a central system). The personalization attributes are defined as: (i) conformity, the importance the UAV gives to the user's suggestions; (ii) sociability, the dispersion among the UAVs; (iii) dedication, the effort put in by the UAV to track the lost target; and (iv) disposition, the UAVs frustration during the negotiation process. These attributes allow the UAVs to behave more efficiently, and not just stick to a centralized control, found in most traditional swarm control approaches.

66.3 Optimization Approaches

This section presents a description of the bio-inspired and probability-based approaches adopted in the previously described MTS techniques.

66.3.1 Ant Colony Optimization (ACO)

Ants possess collective complex social behaviours, which can be implemented to solve different types of combinatorial optimization problems. Specifically, ants tend to find the shortest path from their nest to food by depositing pheromone on the paths, hence creating and following a "pheromone trail'. Since the ants also follow the pheromone trails of other ants, more pheromone on a particular path increases the chance of finding the ideal trail. This ability of ants to find the shortest paths, is the foundation of ACO, a population-based metaheuristic which can be used to estimate solutions for practical optimization problems. This is done by creating a pheromone matrix and then using the matrix to record the paths created by the UAVs. This allows us to find the UAV path with the highest score in the pheromone matrix, which tends to be the suitable trajectory. ACO also allows us to combine the pheromone information with a problem-specific heuristic to guide the UAVs towards better solutions more quickly.

66.3.2 Evolutionary Algorithm (EA)

EAs are a population-based algorithms which incorporate natural selection features into the optimization process. In this algorithm, the population gets better fitness through an iterative process of selection, reproduction and mutation. This algorithm is different from the other approaches, as it randomly samples the solutions that will be crossed. This provides higher chances of obtaining better solutions. It then determines the best solutions that will survive in the following iterations by comparing the new population with the previous one. To handle multiple objective functions and constraints, it contrasts the quality of pairs of solutions based on an overall constraint violation criterion first and on a Pareto comparison of the objective functions second [8]. EA can be applied to maximize the probability of target detection which are searched by the UAVs (whose movement is limited by cardinal directions).

66.3.3 Cross Entropy Approaches (CEO)

CEO uses the cross entropy method, which is a probabilistic-based method used for continuous and combinatorial problems. It is an iterative procedure which can be divided into two phases: (1) using a specified probability function to create random data samples (in our case, trajectories) and (2) updating the parameters of the probability function based on the best data (best current solutions) to construct better samples (solutions) in the succeeding iteration. The CEO approach used in MTS assume the independence among the decision variables of the problem, allowing to generate trajectories that combine the most probable directions to perform at each time step.

66.3.4 Bayesian Optimization Algorithm (BOA)

BOA is a probabilistic optimization algorithm that determines which nodes are the decision variables of the problem, that better fits the group of best solutions of a problem, at each iteration of the Bayesian Network (BN). The BN is then used to sample the set of new solutions that will be evaluated, sorted, and selected as the best set of the following iteration. These two steps are repeated for all iterations as in [1], to determine the relevant relationships among the decision variables, and it returns the best solution found at the end of these iterations.

66.4 Mathematical Evaluation of MTS

This section presents and compares two different models used in the literature to evaluate the expected time of target detection. The two selected models are the ones presented in the introductory work in [1] and the one used in the recent ACO-based work in [11].

In both cases, the unknown target state is modelled with a random variable that has known initial belief and motion probability; the UAV location is known (although not explicitly written in [1]), and the sensor observations can be detection or non-detection measurements, and are related with the target state and UAV location through the sensor likelihood function. All these variables and probability functions used in each paper are summarized in Table 66.2.

66.4.1 Models and MTS Evaluation In [1]

The target initial belief $p(x_0^t)$, target motion model $p(x_k^t|x_{k-1}^t)$ and sensor likelihood $p(z_k|x_{k-1}^t)$ are used in this case to obtain the target belief $p(x_k^t|z_{1:k})$ at any time k given the measurements z_{1_k} up to this time. To do this, it makes use of a recursive bayesian filter that iterates through two steps:

Table 66.2 Variables and probabilities summary

Description	In [1]	In [11]
Time	k	t
UAV indexing	–	u
Target state at given time	x_k^t	v^t
UAV location at given time	–	x_u^t
Sensor observations at given time by given UAV	z_k	z_u^t
Detection observation	D_k	D
Non detection observation	$\overline{D}_k$	$\overline{D}$
Initial belief	$p(x_0^t)$	$P(v_0)$
Target motion probability	$p(x_k^t\|x_{k-1}^t)$	$P(v^t\|v^{t-1})$
Sensor likelihood	$p(z_k\|x_k^t)$	$P(z_u^t\|v^t, x_u^t)$
Probability of not detecting the target	Q_k	$P(z_{1:U}^{1:t} = \overline{D}\|x_{1:U}^{0:t})$
Expected time to detect the target	MTTD	ET

- *Prediction* phase, which is required when the Probability Distribution Function (PDF) of the target state is time dependent since the target is moving or the ambiguity about its position is increasing.
  This step updates the PDF with the latest observations $p(x_{k-1}^t|z_{1:k-1})$ using the Chapman Kolmogorov expression and the target motion model $p(x_k^t|x_{k-1}^t)$:

$$p(x_k^k|z_{1:k-1}) = \int p(x_k^t|x_{k-1}^t) \cdot p(x_{k-1}^t|z_{1:k-1}) \cdot dx_{k-1}^t \tag{66.1}$$

- *Update* phase, which is performed using the Bayesian theorem and the sensor likelihood $p(z_k|x_k^t)$ after the new observation z_k at time k becomes available:

$$p(x_k^t|z_{1:k}) = \frac{1}{K} \cdot p(z_k|x_k^t) \cdot p(x_k^t|z_{1:k-1}) \tag{66.2}$$

  where $K = \int p(z_k|x_k^t) \cdot p(x_k^t|z_{1:k-1}) \cdot dx_k^t$ is a normalization factor to ensure that $p(x_k^t|z_{1:k})$ is a PDF (i.e. $\int p(x_k^t|z_{1:k}) \cdot dx_k^t$=1). Note that for Eq. (2) z_k must equal the value of the observation at time k (i.e. $z_k = D_k$ or $z_k = \overline{D}_k$), and that the likelihood also depends on the UAV location from where the measurement z_k is taken, although it is not implicitly stated in the expression.

To evaluate a given UAV trajectory according to the MTS criterion, the following expressions are used:

- Probability of not detecting $\overline{D}_k$ the target at time step k given previous measurements. $z_{1:k-1}$:

$$r_k = p(\overline{D}_k|z_{1:k-1}) = \int p(\overline{D}_k|x_k^t) \cdot p(x_k^t|z_{1:k-1}) \cdot dx_k^t \tag{66.3}$$

  Note that the expression is similar to the one used to evaluate K, when $z_k = \overline{D}_k$.
- Probability of not detecting the target from UAV trajectory.

$$Q_k = P(\overline{D}_{1:k}) = \prod_{i=1}^{k} r_j \tag{66.4}$$

  Note that the probability P_k of detecting the target from any point of the UAV trajectory, is the complementary of the previous (i.e. $P_k = 1 - Q_k$).
- Probability of detecting the target at time k, not having it detected at $j < k$

$$p_k = (1 - q_k) Q_{k-1} \tag{66.5}$$

- The Mean Time To Detect (MTTD) the target, which is the evaluation criterion itself:

$$\text{MTTD} = \sum_{k=1}^{\infty} k \cdot p_k \tag{66.6}$$

Note that previous equation is a discretized version of the usual expression to obtain the mean value of a variable (i.e. $\text{mean}(t) = \int t \cdot p(t) \cdot dt$), with $p(t) = p_k$

66.4.2 Models and MTS Evaluation In [11]

The target initial belief $P(v_0)$, target motion model $p(v_t|v^{t-1})$ and sensor likelihoods $p(z_u^t|v^t, x_u^t)$ are used in this case to obtain directly an "unnormalized" target belief $\tilde{b}(v^t)$ updated up to time k given only non detection measurements up to this time. To obtain this function it uses the following expression:

$$\tilde{b}(v^t) = \prod_{u=1:U} P(z_u^t = \overline{D}|v^t, x_u^t) \quad \cdot \sum_{v^{t-1} \in G_\Omega} P(v^t|v^{t-1}) \cdot \tilde{b}(v^{t-1}) \tag{66.7}$$

In relation with the previous section's expressions, it is worth to note the following. On one hand, Eq. (66.7) is equivalent to the prediction step stated at Eq. (66.1). Hence, it is used in Eq. (66.7) to redistribute $\tilde{b}(v^{t-1})$ according to the target motion model. On the other hand, Eq. (66.7) is similar to the update step stated at Eq. (66.2). However, Eq. (66.7) includes a product of likelihoods (to account for the measurements of multiple UAVs), lack of the normalization term (required to obtain a proper belief), and assume that the $z_u^t = \overline{D}$ (to be used next to calculate the probability of not detecting the target).

Next, the approach calculates the probability of not detecting the target with Eq. (66.8). This new expression plays the role of Eqs. (66.3) and (66.4) in the previous approach. Moreover, the external summation in Eq. (66.8) stands for the integral in Eq. (66.3), and the product of Eq. (66.4) is not needed since $\tilde{b}(v^t)$ has not been normalized in Eq. (66.7). Hence, this second process is computationally quickly.

$$P(z_{1:U}^{1:t} = \overline{D}|x_{1:U}^{0:t}) = \sum_{v^{t-1} \in G_\Omega} \tilde{b}(v^t) \tag{66.8}$$

The Expected Time (ET) to detect the target is obtained with Eq. (66.9). Although it plays the same role as Eq. (66.6), the ET is calculated taking advantage that the time is a positive variable and hence its mean value can be obtained as $\text{mean}(t) = \int (1 - P(y \leq t)) \cdot dt$, being $(1 - P(y \leq t))$ for this case, the probability of not detecting the target up to time step t (i.e. $1 - P(y \leq t) = P(z_{1:U}^{1:t})$).

$$ET(x_{1:U}^{0:N}) = \sum_{t=1}^{\infty} P(z_{1:U}^{1:t} = \overline{D} | x_{1:U}^{1:t}) \tag{66.9}$$

Finally, it is worth noting the following. Although the summation in Eq. (66.9) and Eq. (66.6) are up to ∞, it has to be truncated for evaluation purposes. This often makes the ET and MTTD expressions underestimate the mean value of the detection time, causing Eq. (66.9) to have smaller truncation errors than Eq. (66.6).

66.5 Conclusion and Future Work

This paper presents a clear comparison between several bio-inspired and probabilistic optimization algorithms used for UAVs path planning in MTS. Two mathematical models (ACO & Bayesian) have been elaborated as well to emphasize the same. On comparing different literature based on MTS, it is seen that ACO is predominantly used to solve the discrete version of the MTS problem (where the UAVs' flight follows straight lines), while EA is mostly used for its continuous counterpart (where the UAVs' flight follows curved smooth trajectories). Besides, many works use idealistic MTS scenarios and only a handful of them successfully take advantage of ad-hoc constructive heuristics to obtain overall good solutions quickly. Hence, our future work will consist of hybridizing new ad-hoc constructive heuristics for MTS with general optimization meta-heuristics, to obtain new approaches for more realistic MTS scenarios.

References

1. Bourgault, F., Furukawa, T., Durrant-Whyte, H.: Optimal search for a lost target in a Bayesian world. In: Field and Service Robotics, pp. 209–222 (2003)
2. Hexmoor, H., Mclaughlan, B., Baker, M.: Swarm control in unmanned aerial vehicles. In: Proceedings of International Conference on Artificial Intelligence (2005)
3. Lanillos, P., Besada-Portas, E., G-Pajares, Ruz, J.: Minimum time search for lost targets using Cross Entropy Optimization. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 602–609 (2012)
4. Lanillos, P., Yañez-Zuluaga, J., Ruz, J., Besada-Portas, E.: A Bayesian approach for constrained multi-agent minimum time search in uncertain dynamic domains. In: Proceedings of the 15th Annual Conference on Genetic and Evolutionary Computation, pp. 391–398 (2013)
5. McGuire, M., Rys, M., Rys, A.: A study of how unmanned aircraft systems can support the Kansas Department of Transportation's efforts to improve efficiency, safety, and cost reduction. Kansas State University Transportation Center, Tech. rep. (2016)
6. Meghjani, M., Manjanna, S., Dudek, G.: Multi-target rendezvous search. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 2596–2603. IEEE, New York (2016)
7. Perez-Carabaza, S., Bermudez-Ortega, J., Besada-Portas, E., Lopez-Orozco, J., de la Cruz, J.: A multi-UAV minimum time search planner based on ACO-R. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 35–42 (2017)

8. Perez-Carabaza, S., Besada-Portas, E., Lopez-Orozco, J., de la Cruz, J.: A real world multi-UAV evolutionary planner for minimum time target detection. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 981–988 (2016)
9. Perez-Carabaza, S., Besada-Portas, E., Lopez-Orozco, J., de la Cruz, J.: Ant colony optimization for multi-UAV minimum time search in uncertain domains. Appl. Soft Comput. **62**, 789–806 (2018)
10. Pérez-Carabaza, S., Besada-Portas, E., López-Orozco, J., Pajares, G.: Minimum time search in real-world scenarios using multiple UAVs with onboard orientable cameras. J. Sens. (2019)
11. Pérez-Carabaza, S., Scherer, J., Rinner, B., López-Orozco, J., Besada-Portas, E.: UAV trajectory optimization for minimum time search with communication constraints and collision avoidance. Eng. Appl. Artif. Intell. **85**, 357–371 (2019)

Chapter 67
A Comparative Study on Breast Cancer Tissues Using Conventional and Modern Machine Learning Models



D. Lakshmi, Srinivas Reddy Gurrela, and Manideep Kuncharam

Abstract Recent advancement in analytical models, availability of hardware resources like GPU, TPU, and cloud infrastructure began to play a pivotal role in healthcare practices and research. It has numerous tools and techniques are available to archive, manage, analyze, and predict large volumes of structured, unstructured, and semi structured data. 'Data Science' plays a vital role in the medical field with the better support for diagnosis, early prediction and cure for the disease. In this research, a comparative study on conventional models and modern methods are carried out using 12 different machine learning models for classification. The response or target variable is the diagnostic observation having two classes viz benign (not cancerous) or malignant (cancerous). The dataset used is The Wisconsin Diagnostic Breast Cancer (WDBC) which is taken from the UCI machine learning repository. The performances of the models are analyzed using the validation matrices, precision, recall, accuracy, and F1 Score. In this research study, we have built the ensemble model using the base estimator as a Random Forest classification model, Principal Component Analysis (PCA), Extra Gradient Boosting (XGBoost) and bagging techniques. Among 11 conventional models, the highest accuracy is obtained for the combined model which is 98.8%. A modern method tSNE is used for dimensionality reduction along with the feature engineering method called k-best method and base model for the classification is Random Forest model is used. This combined approach resulted in the best accuracy 99%. We have obtained a little improved score using PCA and tSNE methods than the reference articles considered in this research.

D. Lakshmi (✉) · M. Kuncharam
B V Raju Institute of Technology, Narsapur, Hyderabad, India
e-mail: lakshmi.d@bvrit.ac.in

S. R. Gurrela
ExcelR, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_67

67.1 Introduction

According to World Health Organization (WHO), breast cancer is the second most common cancer in women worldwide especially in developed countries [2]. With the aid of research and development support, the possibility of early detection paves the way to reduce the death rate. Machine learning algorithms are extremely useful in early prediction in order to cure the disease and save the patient.

The main focus of this research study is to classify the patient tumor measurements and medical observations into benign (not a cancer) or malignant (cancerous) using 10 conventional models and one modern method named tSNE. Whereas, 'Benign' is a medical term that signifies that the tumor cells that are not cancerous and will not spread to other areas of the body. 'Malignant' is a medical term that signifies that the tumor is made of cancer cells, and it will affect nearby tissues.

Previous researches using the Wisconsin Diagnostic Breast Cancer (WDBC) [1] dataset have shown high-quality result using machine learning models [2–4, 6–10]. The research study, published in November 2019 by [2] has shown 98.24% accuracy using Genetic programming approach with Adaboost classifier. Similarly, [4] has shown the results with 95.0791% accuracy for the Random Forest classifier. In our research study obtains 98.83% accuracy with convention models and 99% with modern method when compared to the highest accuracy obtained from the research studies [2] and [3].

67.2 Conventional Classification Models

In this section, a comparative study on 11 conventional classification models are discussed based on model validation metrics.

67.2.1 Data Description

There are 30 diagnostic measurements or features available in the original dataset. Among these, 29 features are of real number type that are considered as independent variables (X-Variable). The 'Diagnosis' variable is considered as the dependent variable (Y-Variable/Predictive Variable). There are 212 malignant and 357 benign observations are considered for the classification.

67.2.2 Methodology

In this research study, there are 11 classification models which are built to classify the dataset. It follows the various data analytics stages such as data collection, data preprocessing, data transformation, feature selection, and model validation and tuning the hyperparameters in order to increase the accuracy rate. The proposed classification model is used in the WDBC dataset following the sequence of steps like data standardization, dimensionality reduction, hyperparameter tuning, and development of classification model with boosting, bagging, and model validation.

67.2.3 Data Standardization

Data standardization is the basic requirement for most of the machine learning algorithms and also to obtain optimal performance. Data standardization is done on this dataset using 'Min–Max-Normalization'. This method transforms all the features onto the unit scale with the mean as 0 and variance as 1.

67.2.4 Model Development

For all the classification models 70% of the dataset is considered in the training phase and 30% in the testing phase. For the training phase, tenfold cross validation is considered for the classification models.

67.2.5 Model Improvement Techniques

Hyper parameter tuning is the process of optimizing the model parameters to get the best accuracies on train, test, and validation data sets. In this research study, there are three strategies are considered for the model tuning or hyperparameter tuning, such as Grid Search, Random Search, and k-Fold validation techniques.

Grid search is a conventional way to find optimal parameters for a model and it needs search space to be given and model would be built on all the values of the search space and accuracies will be computed. Combinations of hyper parameters will be selected based on best accuracy.

Random search differs from grid search mainly in that it searches the specified subset of hyper parameters randomly instead of exhaustively. The major benefit of this method is with decreased processing time.

In the process of model validation, data set have been divided into 10 folds. There are 9 folds are considered for the training phase and tested on onefold for each

iteration. This process was repeated 10 times with different folds in train and test data.

67.3 Results and Discussion

The model validation measures used in classification algorithms are confusion matrix, precision, recall, F1-Score, accuracy shown in Fig. 67.1.

67.3.1 Comparative Study on 10 Classification Models

In the research study [5], there is an imbalance in the predictions with the Support Vector Classifier model (Benign 78.72%, Malignant 0%, and Average 51.10%), whereas, in our research study imbalances have been minimized. And we have obtained 99% for Benign and 93% for Malignant and an average of 95.10%. Although all the 11 classification models are chosen randomly, first model is started with Dummy Classifier as a baseline classifier. All the classification models which are implemented in this study resulted with more than 90% accuracy. However, for the ensemble model Random Forest algorithm is chosen as a base algorithm and combined with PCA, and AdaBoost techniques in order to improve the accuracy. As

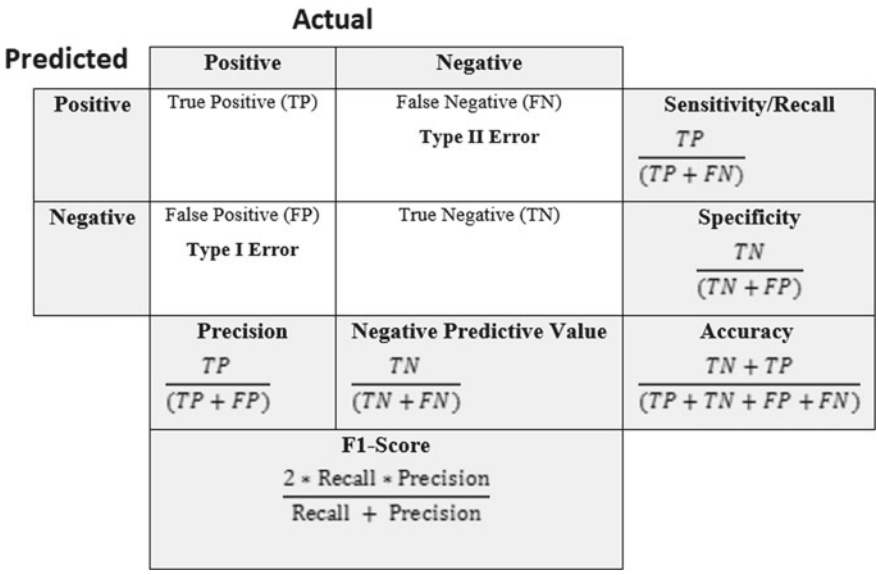


Fig. 67.1 Confusion matrix and validation metrics used in classification models

shown in the Table 67.1, Random Forest classifier resulted with 100% as a highest precision score, Logistic Regression, and Decision Tree Classifier resulted in 98.85% as a highest recall score, Random Forest Classifier using AdaBoost and PCA resulted in 98.74% as a highest F1-Score and 98.83% as a highest accuracy.

Table 67.1 Comparison of 10 classifier models precision, recall, f1-score, and accuracy

S. No.	Model	Precision score (%)	Recall score (%)	F1-score (%)	Accuracy score (%)
1	Logistic regression	93.47	*98.85*	96.08	95.10
2	Linear discriminant analysis (LDA)	*100*	87.35	93.22	95.32
3	Support vector classifier	93.47	98.85	96.08	95.10
4	Gaussian Naïve Bayes	94.25	94.25	94.25	93.00
5	Decision tree classifier	96.20	87.35	91.56	90.20
6	K nearest neighbourhood classifier	98.76	91.95	95.23	94.40
7	Random forest classifier using AdaBoost and PCA	99.16	98.33	*98.74*	*98.83 (0.5)*
8	Gradient boosting using decision tree as a base algorithm	95.40	95.40	95.40	94.40
9	Extra gradient boost using decision tree as a base algorithm	98.30	98.30	98.30	98.00
10	Random forest classifier	96.51	95.40	95.95	95.10
11	Bagging classifier	91.5	97.9	95.8	95.5

67.4 Modern Method Using tSNE

The conventional model named Principal Component Analysis (PCA) is linear transformation technique and this model would lead to poor performance especially when dealing with non-linear manifold structures.

The t-Distributed Stochastic Neighbor Embedding (t-SNE) model has emerged as an admired model for data visualization. It also helps in dimensionality reduction and also for feature extraction. This model has major three steps namely, similarity computation in higher dimensional space, similarity computation in lower dimensional space and minimizes the difference between these two based on conditional probabilities or similarity score in high dimensional and low-dimensional space.

The modern method for dimensionality reduction using t-Distributed Stochastic Neighbour Embedding (t-SNE) and feature engineering method select k-best method, which works based on chi-square method. We have got the best features among all the features ($n = 28$). The base model of the classification is Random Forest Algorithm, which is a bagging method.

SelectKbest method is used for the feature engineering. It is univariate feature selection method. It works by selecting the best features based on univariate statistical tests. This method can be seen as a preprocessing step to an estimator. Scikit-learn exposes feature selection routines as objects that implement the transform method. SelectKBest removes all but the highest scoring features.

Random Forest base classifier is used with 1000 trees and max_features = 5 among 28 features given in the original dataset. The dataset is split with (80:20) for the train and test phase respectively. The tSNE model validation results are shown in the Table 67.2.

A modern method tSNE is used for dimensionality reduction along with the feature engineering method called k-best method and base model for the classification is Random Forest model. This combined approach resulted with the best accuracy 99%.

Among three dimensionality reduction techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and t-Distributed Stochastic Neighbour Embedding (tSNE), tSNE has resulted with the highest accuracy 99%.

Table 67.2 tSNE model validation matrices

S. No.	tSNE, bagging and random forest classification model	Precision score (%)	Recall score (%)	F1-score (%)	Accuracy score (%)
1	Training phase	98.59	97.61	97.9	99.05
2	Testing phase	95.03	97.22	96.47	99.03

67.5 Conclusion

This research study is conducted in order to find out the best suitable ensemble model on WDBC dataset with the high accuracy of class label prediction. All the 11 classification models are implemented using Python libraries. The ensemble model Random Forest algorithm is chosen as a base algorithm and combined with PCA, AdaBoost, and Bagging are found to be superior among all the 10 conventional models with the highest accuracy of 98.8%. A modern strategy tSNE is utilized for dimensionality reduction along with the feature engineering method called k-best method and base model for the classification is Random Forest model. This ensemble model resulted in the greatest accuracy with 99%. The Python code for all the 11 models are available in the GitHub for the references.

References

1. Wisconsin Diagnostic Breast Cancer (WDBC) dataset: https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29
2. Dhahri, H., Maghayreh, E.A., Mahmood, A., Elkilani, W., Nagi, M.F.: Automated breast cancer diagnosis based on machine learning algorithms. J. Healthc. Eng. (2019)
3. Nidhi, M.K., Makkar, S.: Classification of breast cancer tissues using decision tree algorithms. Int. J. Res. Eng. Appl. Manage. **04**(05) (2018). https://doi.org/10.18231/2454-9150.2018.0636
4. Salama, G., Abdelhalim, M.B., Zeid, M.: Experimental comparison of classifiers for breast cancer diagnosis, 180–185 (2012). https://doi.org/10.1109/ICCES.2012.6408508
5. Ferri, C., Hernández-Orallo, J., Modroiu, R.: An experimental comparison of performance measures for classification. Pattern Recogn. Lett. **30**(1), 27–38 (2009). https://doi.org/10.1016/j.patrec.2008.08.010
6. Christobel, A., Sivaprakasam, Y.: An empirical comparison of data mining classification methods. Int. J. Comput. Inf. Syst. **3**(2) (2011)
7. Lavanya, D.: Ensemble decision tree classifier for breast cancer data. Int. J. Inf. Technol. Converge. Serv. **2**(1), 17–24 (2012)
8. Ster, B., Dobnikar, A.: Neural networks in medical diagnosis: comparison with other methods. In: Proceedings of the International Conference on Engineering Applications of Neural Networks, pp. 427–430 (1996)
9. Joachims, T.: Transductive inference for text classification using support vector machines. In: Proceedings of International Conference Machine Learning. Slovenia (1999)
10. Abonyi, J., Szeifert, F.: Supervised fuzzy clustering for the identification of fuzzy classifiers. Pattern Recogn. Lett. **14**(24), 2195–2207 (2003)

Chapter 68
Clinical Decision Support System for Knee Injuries Treatment Using Multi-Agent System



Naveen Dalal and Indu Chhabra

Abstract Diagnosis of sports injuries is a very critical process and its performance depends on the recognition of the relevant symptoms. In this paper, a Multi-Agent-based knee injury detection and diagnosis scheme (MFZS) is introduced that applies fuzzy rules over input symptoms and recommends the relevant treatments. Its performance is compared with traditional fuzzy system (TDFZS) under the constraints of detection accuracy and sensitivity etc. Chi-square and Fisher Exact test is also performed to verify the significance of the outcomes of both schemes.

68.1 Introduction

Clinical decision support systems can be enhanced using the applications of artificial intelligence. Expert systems can improve the overall diagnosis process and sports team can predict the injuries and risks related to a specific sports activity/exercise. Figure 68.1 shows the interaction of different entities with expert system.

Support of medical practitioners is used to design and refine the expert system. Performance and injury risk management may be initiated by sports team whereas injured team member may acquire help regarding diagnosis from expert system. Communication between expert system and these entities suffers from the following barriers and challenges:

Barriers for Medical Practitioners

Designing an expert system requires experience man power as well as highly accurate input data sets to produce rules and response. Sports injury data collection and its

N. Dalal (✉)
Goswami Ganesh Dutta Sanatan Dharam College, Chandigarh, India

I. Chhabra
DCSA, Panjab University, Chandigarh, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_68

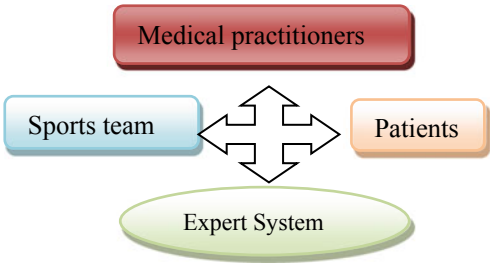


Fig. 68.1 Entities interaction with expert system

Table 68.1 Challenges for sports team/patients

Challenges for sports team	Challenges for patients
Sports teams may engage in different activities in a specific sports type. Occurrence of an injury depends on the following factors: experience of coach/team member, confidence level, type and frequency of exercises involved in a given sports. All above factors may increase or decrease the risk factor associated with sports injuries. Using an expert system, participants can predict the injuries during an activity as well as medical practitioners can predict the treatment type with respect to injury type	Expert system uses its knowledge base as per the input symptoms provided by current patient and produces a treatment plan. Error prone input may alter the treatment type and patient recovery will suffer due this mismatch. Expert system may produce accurate treatment plan, however supervision of a medical practitioner is still required to crosscheck the treatment recommendations [1–5]

sorting is a very complex task thus may reduce the overall system performance and diagnosis accuracy. Diagnosis prediction may suffer due to obsolete rules and datasets, so medical practitioners should update these input parameters as per current requirements. Large-scale datasets and rules can increase the response time. So there is a need to optimize the rules of an expert system (Table 68.1).

68.2 Requirements of Intelligent Agents for Injury Detection and Diagnosis

Following are the requirements of intelligent agents for injury detection and diagnosis:

In case of single diagnosis system, treatment recommendations are generated as per the input systems and it will not generate no response, if it fails to recognize the input whereas in case of multi agent system, if a diagnosis agent fails to identifies the input symptoms, then request is forwarded to another agent and so on. If all agents fail to resolve the category of treatment, only then treatment for input injury is refused. Data set can be subdivided and assign to each agent, in order to produce quick response. It will also minimize the number of searches to match the input

symptoms with available treatments. As per requirements, only specific agent can be invoked for treatment and there is no need to engage all agents in diagnosis process, in this case, resource utilization can be optimized.

68.3 Contribution of Fuzzy Logic in Medical Domain

Kasbe et al. [6] introduced a solution to diagnosis heart disease. Centroid method is used with fuzzy logic for defuzzification process. A dataset is used to produce multiple inputs and ranges. Finally, fuzzy rules are used over output variations (0–5) to determine the disease conditions. Experimental outcomes show its performance in terms of higher accuracy as compared to traditional diagnosis systems. Shaik et al. [7] introduced a fuzzy logic based anemia diagnosis system that takes blood parameters as input and applies fuzzy rules to generate output. Experimental results show that it can identify the various types of disease and can enforce the rules accordingly. Experimental results indicate that it can detect/diagnosis the disease in real-time environment. Thamaraimanalan et al. [8] developed a solution for handheld medical devices that can be used for heart patients. It uses multiple digital signals as input and then fuzzy rules are used to analyze the results. Analysis shows that it can monitor the heart rate of patients with higher accuracy as well as it consumes optimal energy. Farzandipour et al. [9] did the investigation of fuzzy logic based diagnosis solutions and found that their accuracy and outcomes can be refined using the various datasets. Study shows that clinical knowledgebase can be utilized to develop new fuzzy rules and that can be utilized by the healthcare industry for decision making. Akinnuwesi et al. [10] developed a fuzzy logic based expert system for rheumatology. It considers various aspects (symptom and degree of uncertainty/associated risk and causal factors) for diagnosis and decision making. Experimental results show that it can enhance the overall efficiency of healthcare practitioners/diagnosis accuracy etc. as compared to traditional diagnosis solutions. Bautista et al. [11] investigated the relationship between technology, disease, diagnosis and decision makers. Study shows that mapping of all above factors can be used to build highly accurate algorithms (fuzzy logic based) for medical expert system for healthcare industry. Patil et al. [12] explore the various artificial intelligence (AI) based tools (neural networks/fuzzy logic/machine learning) that can be used for healthcare services. Study shows that integration of these tools with BigData can be used to produce a large scale knowledgebase for artificial decision support system for healthcare industry. Singh et al. [13] investigated the contribution of AI-based methods that can be used for the diagnosis of injuries related to neuro system. It considers various factors (neurological pathways/cognition/neurotransmission/disability/behavior disorders) that can be utilized to form the predicates for machine learning algorithms. Analysis shows that AI-based healthcare solutions can maximize the accuracy of diagnosis process. Medeiros et al. [14] presented an interference method for healthcare that uses fuzzy logic to map the risk assessment of patients with decision support system, in order to improve the overall diagnosis process. Analysis shows that mapping of

above-discussed parameters can be utilized to develop training datasets for practitioners. Kadhim et al. [15] merged the neural network with fuzzy logic and developed a framework to diagnose the back pain. It uses different health parameters to perform fuzzification/defuzzification. Experimental results show its performance in terms of higher efficiency/accuracy of diagnosis process. Alsmadi [16] used the combination of neutrosophic with fuzzy logic to detect the lesions. A metric is used to define the final output and experimental results indicate that it outperform in terms of accuracy, efficiency as compared to traditional lesions detection schemes. Ahmadi et al. [17] investigated various decision support methods that use fuzzy interference/neural network/decision tree/rational mapping/weighted rules for diagnosis of different diseases (Joint pain/Arthritis/Cancer/Cardiac/Malaria/Hypovolemi). Analysis displays that efficiency and accuracy of diagnosis process can be improved using these methods and the performance of the developed schemes depends over diverse parameters, i.e., sample size/interval/rules etc. Mardani et al. [18] investigated the relationship between technology, domain, healthcare services, environment control, medical waste, risk, patients, practitioners, and diagnosis process, etc. Study shows that there are various AI-based solution that can be used boost the decision making. Analysis identifies some open issues related to selection of fuzzy sets, aggregation operators, rules complexity and standardization communication interface. Mohandes et al. [19] developed a health risk assessment framework for construction industry. It considers the various factors associated with current health, treatment and risks, fuzzy rules, etc. Experimental results show that it can assists the practitioners to track the health issues/diagnosis process and it can be further extended to analyze the relationship between risks and accidental hazards, in order to reduce their probability factor. Lohrmann et al. [20] developed a customized fuzzy feature selection method for diagnosis. It considers various parameters i.e. original features/scale factors/multiple datasets etc. and performs classification to reduce the distance between different features. Similar features are reduced to their minimal value to produce unique feature sets and final outcomes are considered for diagnosis process. Experimental results show its performance in terms of diagnosis accuracy as compared to traditional feature selection/filtering methods.

68.4 Multi-Agent Based Diagnosis

In case of sports, knee injury is a common risk to team that can cause serious health issues. This paper introduces a multi-agent based diagnosis scheme for these injuries. Figure 68.2 shows the involvement of the different agents for injury detection and diagnosis at college/university/hospital level.

User Agent (UA): User Agent gets injury state of the patients and sends it to Diagnoses Agent (DA) for analysis, diagnose, and treatment. Further UA shows the results and treatment plan generated by the DA to the user. It will also respond to the query from the Diagnosis Agent.

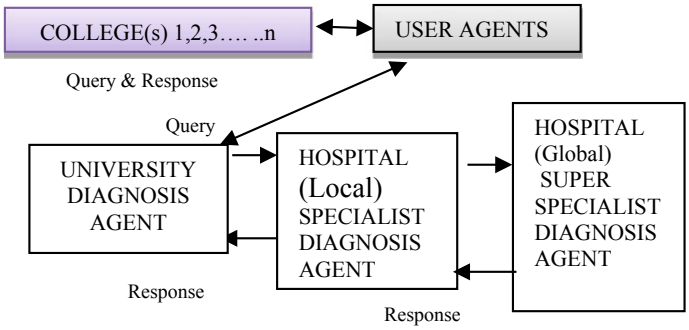


Fig. 68.2 Multi-agents for diagnosis process agents description

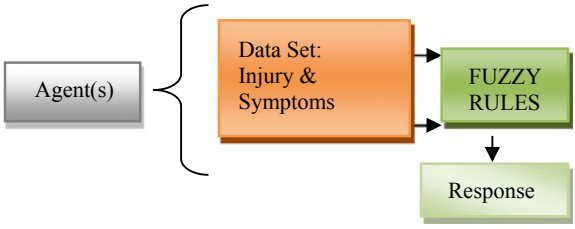
UA deals with the data of coaches/doctors, patients, injury details.

Diagnosis Agent (DA): This agent is fully automated and no human intervention is required to diagnosis the sports injury. Its functions are to analysis about patient's state with integration information. The DA will react and ask for more information if necessary and generate treatment plan accordingly. In complicated cases if it needs to consult Specialist Diagnosis Agent (SDA) it can consult. There is one DA at each level i.e. university, hospital in same region, hospital in same national zone (north, south, east, and west). At each level the specialty of DA increases. DA is further categorized as SDA and Super Specialist Diagnosis Agent (SSDA). DA will also save the user data about their personal information, treatment plan for further reference.

Fuzzy Rules for Multi-Agents

Figure 68.3 shows the Fuzzy rules for Multi-agents, in which all agents use their injury and symptom datasets as input and fuzzy rules are applied to produce the final response for all agents are applicable and following is the basic procedure used for query and response. We have used Mamdani Inference method.

Fig. 68.3 Fuzzy rules for multi-agents



```
RULE 1: IF (kneepain IS low) OR (grade IS gradeone) THEN (treatmentplan IS selfcare);
RULE 4: IF (kneepain IS low) OR (grade IS gradeone) OR (symptoms IS bleeding) THEN (tplan
IS Coldtherapy);
RULE 43: IF (kneepain IS low) OR (grade IS gradeone) OR (symptoms IS instability) THEN
(tplan IS Coldtherapy);
RULE 69: IF (kneepain IS high) AND (grade IS gradethree) AND (symptoms IS
Inabilitytofullystraightenknee) THEN (tplan IS rehabilitation);
```

Step 1: Fuzzification defines membership functions. For instance, we defined knee pain into four linguistic terms, namely symptoms, knee pain, knee strain, and grade

Step 2: Designing the knowledge base by fuzzy rules using an if–then format. These rules are defined using linguistic terms such as:
IF (knee pain IS low) OR (grade IS grade one) THEN (treatment plan IS self-care).
For each linguistic term, a corresponding fuzzy set can either be designed with the help of expert(s) or derived from existing data set.

Step 3: Rule Evaluation is done using input that cause various rules to fire with varying strengths. For instance, many knowledge based rules are fired out for the symptoms, level of knee pain, grade of injury and knee strain. This leads to decisions regarding potential treatment(s), which can then be administered. The outcome of the rules are evaluated to reach to the conclusion.

Step 4: Defuzzification: finding the crisp outcome for output variables is important in this proposed system. As its needs to inform the doctor which treatment plan should be performed, e.g., self-care, cold therapy, surgery, etc.

Algorithm

UA: User Agent, DA: Diagnosis Agent, SDA: Special Diagnosis Agent, SSDA: Super Special Diagnosis Agent, IJS: Injury DataSet, Sym: Symptom DataSet, Usr: Users, AGUsr: User Agent, TP: Treatment Plan.

Initialize (DA, Univ.), IJS {ij1, ij2 … ijn}, (SDA, Hospital (Local)), (SSDA, Hospital (Global)), (IJS, DA), (IJS, SDA), (IJS, SSDA).

//Query (Sender, Receiver), UA sends a query to DA	//DA sends a query to SDA
Query(*AGUsr*-> *Sym, DA)* **If FIND** (Query-> *Sym*, *IJS, FZR*) = = TRUE Generate_RESPONSE (*TP*, **Query**-> *Sym*) **Else** Generate_RESPONSE (REFER_TO(SDA),**Query**-> *Sym*) **End If**	**Query**(*DA, SDA)* **If FIND(**DA.**Query**-> *Sym*, *IJS, FZR*) = = TRUE Generate_RESPONSE(*TP*, DA. **Query**-> *Sym*) **Else** Generate_RESPONSE(REFER_TO(SSDA), DA. **Query**-> *Sym*) **End If**
//SDA sends a query to SSDA	// Response (Sender, Receiver), if invalid symptoms provided by other agents
Query(*SDA,S SDA)* **If FIND(S**DA.**Query**-> *Sym*, *IJS, FZR*) = = TRUE Generate_RESPONSE(*TP*,SDA.**Query**-> *Sym*) **Else** Generate_RESPONSE(NOT_FOUND,SDA.**Query**-> *Sym*) **End If**	**If (**response = = NOT_FOUND) ADVISE = "No Rule found for diagnosis for current symptom(s), Please provide additional information about the injury" Generate_ADVISE (ADVISE)) **End if**

The above algorithm considers **Injury type**: Muscle injuries **Categories** Quads strain **Symptom**: Sitting Problem, Pain, Climbing Stairs, Sensation, Popping Sound, Bruising, Difficulty Walking. **Treatment**: Rest, Physiotherapy, Surgery. **Symptom are classified into**: LOW/MEDIUM/HIGH **Injury Grade is**: ONE/TWO/THREE.

68.5 Results and Performance Analysis

As per Table 68.2, JADE 4.5 was used as design tool for experimental purposes with jre-7 environment over windows platform with different agents (User Agent/Diagnosis Agent/Special Diagnosis Agent/Super Special Diagnosis Agent).

Registration processes of different diagnosis agent is shown in Fig. 68.4 As per the injury type, different systems can be added, and later on all these are used for by diagnosis process. Figure 68.5 shows that diagnosis agent will generate the response only as per the dataset otherwise request of user agent is refused and user agent will wait for other diagnosis agent for diagnosis. Figure 68.5 shows the message exchange between different agents. It can be observed that request of user agent is sent to a diagnosis agent that can accept or refused the incoming request as per the availability of current injury data set. In case of refusal, another diagnosis agent may accept this request, if it supports the diagnosis for current request.

The Multi-Agent graphical user interface (GUI) has been designed to receive symptoms from the doctor and to display possible treatment plan(s) for knee injury. Following are the result given by Multi-Agent System (MFZS) and Fuzzy System (TDFZS) for the calf strains injury Table 68.3.

Table 68.2 Configuration for analysis

Parameters	Description
Design tool	Java agent and development environment (JADE 4.5)
Environment	Java run time environment-7
Platform	Windows 10 (64 Bit)
Agents	User agent, diagnosis agent, special diagnosis agent, super special diagnosis agent
Scenarios	Performance analysis with: (a) traditional Fuzzy Logic Scheme (TFLS), (b) multi-agent based fuzzy logic scheme (MAFLS)



Fig. 68.4 Registration process of Agents (DA/UA)

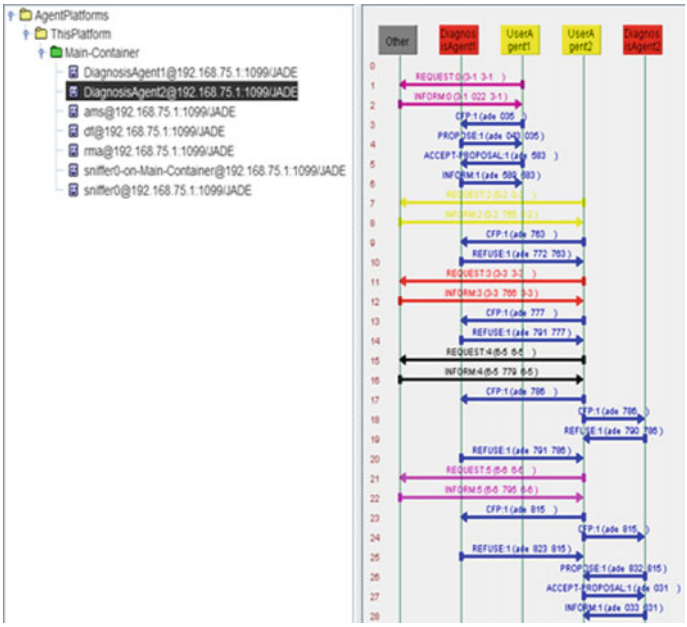


Fig. 68.5 Dynamic response of diagnosis agents

Table 68.3 Treatment recommendations by multi-agent system (MFZS) and fuzzy system (TDFZS)

Muscle injuries								
Category	Non surgery					Surgery		
		Elevation	Medication	Rest	Sports massage	Surgery	Unknown	Total
Calf strains	TDFZS	9	10	11	10	21	0	61
	MFZS	10	11	20	10	10	0	61

Table 68.4 Chi-square test: calf strains

Treatment			
	Non-surgery	Surgery	Total
TDFZS	40 (45.5) [0.66]	21 (15.5) [1.95]	61
MFZS	51 (45.5) [0.66]	10 (15.5) [1.95]	61
Total	91	31	122 (grand total)

68.6 Statistical Analysis

For about 1000 sample cases tested for different knee injuries the responses of the system and the doctor were collected and tabulated (Table 68.3) for calf strain. To verify the homogeneity of the two observations, a Chi-square test and Fisher Exact test of homogeneity of variance was applied on different knee injuries Table 68.3. show data about Calf strain:

a. **Data**: See Table 68.3.
b. Assumption:
c. It is assumed that we have a simple random sample from each one of the two populations of interest
   H0: The two populations are homogeneous with respect to treatments
   HA: The two populations are not homogeneous with respect to treatments Let $\alpha = 0.05$
d. Hypothesis:
   H0: The two populations are homogeneous with respect to treatments.
   HA: The two populations are not homogeneous with respect to treatments.
e. **Statistical Formula**:
   Sum all the values for (O–E) 2 /E.
f. **Calculation of test statistics**: MS-Excel was used for calculating the values.

Table 68.4 Specify the result of Chi-square test for surgery and non-surgery category for Calf Strain. The value calculated for Calf Strain is 5.2329 and *p*-value is 0.022164 and results are significant at $p < 0.05$. Chi-square test with Yates correction is 4.3247 with *p* value 0.037563 and results are significant at $p < 0.05$. The Fisher exact test statistic value is 0.0366 and results are significant at $p < 0.05$. In case of chi square and fisher exact test, significance level floats between <0.1 and <1.0 and for most injury sets, it is significant at <0.5 using both schemes.

68.7 Conclusion

In this paper, issues and solutions related to contribution of expert systems in medical domain, role of different stack holders in the development these systems, have been discussed and a new fuzzy logic based multi agent scheme was introduced to guide the diagnosis process of sports knee injuries. Its performance is also compared with the existing fuzzy logic scheme under the constraints of accuracy/sensitivity. Performance analysis shows that each scheme has its own response for different set of knee injuries. And in some cases, both have recommended same treatment and in some cases, both produced different results that affect their accuracy/sensitivity levels which are directly proportional to each other. Four types of knee injury types have been used for experiments and each type has further subdivided into multiple categories. In case of Muscle Injuries, both schemes have 100% accuracy with highest

sensitivity level (1) for the treatment recommendation for Torn Quad injury whereas in other categories (Calf Strains/Pulled Quad/Quads Strain), (Figs. 68.6 and 68.7) it is slightly declined for TDFZS as compared to MFZS. In case of Knee cap injury type, both schemes have 100% accuracy with the highest sensitivity level for the category: Dislocated Patella Medial Reefing MPFL Repair MPFL Reconstruction, Tibial Tubercle Transfer, and it is slightly decreased for Patella fractures/Medial Imbrication categories. In case of ligament Injuries, for categories: Anterior cruciate ligament, medial collateral ligament, PCL Knee Injury, Ligament Sprains, etc. accuracy and sensitivity level of TDFZS varies whereas it remains almost constant for MFZS. In case of Cartilage Injuries type, for Category: Meniscus Tears, MFZS has the higher accuracy/sensitivity level of treatment recommendation as compared to TDFZS. It can be observed that in case of each injury category, accuracy and sensitivity level of both schemes varies and in few cases, it is approximately constant. As per the above discussion, it can be concluded that MFZS is more sensitive with respect to input symptoms as compared to TDFZS and it can recommend the accurate treatment types to the patients. Current scope of the work considers the limited sports knee injuries types. In future, MFZS may be utilized to diagnose the other injury types as well. Its performance will be optimized using machine learning algorithms.

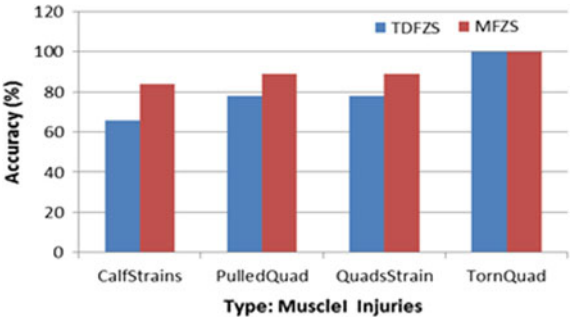


Fig. 68.6 Accuracy of treatment recommendations for muscle injuries

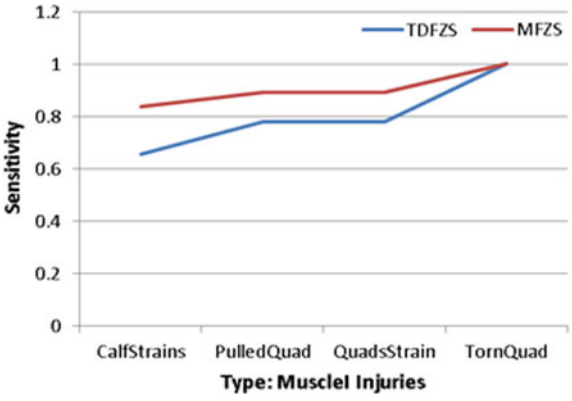


Fig. 68.7 Sensitivity of treatment recommendations for muscle injuries

References

1. Sloane, E.B., Silva, R.J.: Clinical Engineering Handbook, Elsevier 2nd Edition Chapter 83, pp. 556–56 (2020)
2. Fister, I., Ljubič, K., Suganthan, P.N., Perc, M., Fistera, F.: Computational intelligence in sports: challenges and opportunities within a new research domain. Appl. Math. Comput. **262**, Elsevier, 178–186 (2015)
3. Aframian, A., Iranpour, F., Cobb, J.: Medical devices and artificial intelligence. Artif. Intell. Healthc. Elsevier, 163–177 (2020)
4. Valle, X., Geli, E.A.L., Tol, J., Hamilton, B., Pruna, R., Til, L., Gutierrez, J.A., Alomar, X., Balius, R., Malliaropoulos, N., Monllau, J.C., Whiteley, R., Witvrouw, E., Samuelsson, K., Rodas, G.: Muscle injuries in sports: a new evidence-informed and expert consensus-based classification with clinical application. Sports Med. **47**, Elsevier, 1241–1253 (2019)
5. Witkowski, E., Ward, T.: Artificial intelligence assisted surgery. Artif. Intell. Healthc. Elsevier, 179–202 (2020)
6. Kasbe, T., Pippal, R.S.: Design of heart disease diagnosis system using fuzzy logic. In: International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS). IEEE, pp. 3183–3187 (2017)
7. Shaik, M.F., Subashini, M.M.: Anemia diagnosis by fuzzy logic using LabVIEW. In: International Conference on Intelligent Computing and Control (I2C2), pp. 1–5. Coimbatore, IEEE (2017)
8. Thamaraimanalan, Sampath, T.P.: A low power fuzzy logic based variable resolution ADC for wireless ECG monitoring system. Cognit. Syst. Res. **57**, Elsevier, 236–245 (2019)
9. Farzandipour, M., Nabovati, E., Saeedi, S., Fakharia, E.: Fuzzy decision support systems to diagnose musculoskeletal disorders: a systematic literature review. Comput. Methods Programs Biomed. **163**, Elsevier, 101–109 (2018)
10. Akinnuwesi, B.A., Adegbite, B.A., Adelowo, F., Edomwonyi, U.I., Amumeji, O.T.: Decision support system for diagnosing rheumatic-musculoskeletal disease using fuzzy cognitive map technique. Inform. Med. Unlocked **18**, Elsevier, 1–19 (2020)
11. Bautista, J.A.R., Zavala, A.H., Cárdenas, S.L.C., Ruela, J.A.H.: Review on plantar data analysis for disease diagnosis. Biocybern. Biomed. Eng. **38**(2), Elsevier, 342–361 (2018)
12. Patil, S., Patil, K.R., Patil, C.R., Patil, S.S.: Performance overview of an artificial intelligence in biomedics: a systematic approach. Int. J. Inf. Technol. Springer, 1–11 (2018)
13. Singh, R.O.B., Vishweswaraiah, S., Er, A., Aydas, B., Turkoglu, O., Taskin, B.D., Duman, M., Yilmaz, D., Radhakrishna, U.: Artificial Intelligence and the detection of pediatric concussion using epigenomic analysis. Brain Res. **1726**, Elsevier, 1–22 (2020)
14. Medeiros, I.B.D., Machado, M.A.S., Damasceno, W.J., Caldeira, A.M., Filho, J.B.D.S.: A fuzzy inference system to support medical diagnosis in real time. Proc. Comput. Sci. **122**, Elsevier, 167–173 (2017)
15. Kadhim, M.A.: FNDSB: a fuzzy-neuro decision support system for back pain diagnosis. Cognit. Syst. Res. **52**, Elsevier, 691–700 (2018)
16. Alsmadi, M.K.: A hybrid fuzzy C-means and neutrosophic for jaw lesions segmentation. Ain Shams Eng. J. **9**(4), Elsevier, 697–706 (2018)
17. Ahmadi, H., Gholamzadeh, Shahmoradi, M.L., Nilashi, M., Rashvan, P.: Diseases diagnosis using fuzzy logic methods: a systematic and meta-analysis review. Comput. Methods Programs Biomed. **161**, Elsevier, 145–172 (2018)
18. Mardani, Hooker, R.E., Ozkul, S., Yifan, S., Fei, G.C.: Application of decision making and fuzzy sets theory to evaluate the healthcare and medical problems: a review of three decades of research with recent development. Expert Syst. Appl. **13715**, Elsevier, 202–231 (2019)

19. Mohandes, S.R., Zhang, X.: Towards the development of a comprehensive hybrid fuzzy-based occupational risk assessment model for construction workers. Safety Sci. **115**, Elsevier, 294–309 (2019)
20. Lohrmann, C., Luukka, P., Sabuka, M.J., Kauranne, T.: A combination of fuzzy similarity measures and fuzzy entropy measures for supervised feature selection. Expert Syst. Appl. **11015**, Elsevier, 216–236 (2018)

Chapter 69
Predictive Loan Approval Model Using Logistic Regression



Aiswarya Priyadarsini Behera, Siddharth Swarup Rautaray, Manjusha Pandey, and Mahendra Kumar Gourisaria

Abstract Loan disbursement is one of the fundamental aspects of all banks and financial companies. The profit that comes from loan disbursement is one of the major source of bank's assets. Many applicants apply for a loan in any bank or financial company out of which choosing the appropriate costumer is a herculean task. Though banks nowadays have adopted a more rigorous process for validation and corroboration of a costumer, it still does not guarantee that whichever costumer is chosen has all the right and original required documents. The primary objective of this validation is to choose appropriate costumers, so that their assets can be in good hands and help people who are in need. In this work, we propose a loan recommendation system that will provide an immediate and simple way to choose the right applicant based on validation of features. The aim of this paper is to predict a model for loan disbursement by using regression model. Some weights are given to each feature based on the priority of the bank. This system can predict the feasibility of an applicant for availing the loan. It is beneficial for both employees of the bank and the probable candidates for loan disbursement. Here, we will take the important features from loan dataset. The weight of each appropriate feature can be calculated in loan prediction system. For new test data, same attributes can be processed with their respective weights. A particular ceiling time can be defined to allow the applicants to check the status of their application. A jumping mechanism can be added in the prediction system to allow disbursement of loan on a priority basis. However, one disadvantage of the system being in real-life scenario is that the recommendation system may be biased toward one dominant feature or attribute.

A. P. Behera (✉) · S. S. Rautaray · M. Pandey · M. K. Gourisaria
School of Computer Engineering, KIIT (Deemed to be University), Bhubaneswar, India

S. S. Rautaray
e-mail: siddharthfcs@kiit.ac.in

M. Pandey
e-mail: manjushafcs@kiit.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_69

69.1 Introduction

The significance of data analytics in the bank or any financial services sector has been realized at a huge scale. We know that in this technical and advance world, machine is doing various jobs of human and saves more time. Before the AI was introduced, we consumed a lot of time for verification and validation to identify applicant's eligibility. Here, we used logical regression which is also known as logit regression model as machine learning tool for creating a probabilistic and predictive approach to a loan approval prediction model. In regards with this paper, one is the concept in feature selection which is the important attribute to train the model and also apply to a domain of machine learning which is exploited. Advanced analytics assets on a large-scale developing and investing roles are critical to analyze. The applicant revolution enabling clearly defined data ownership and maintenance of high-quality data. A simple linear regression can be replaced by a logistic regression when the input variables are mapped continuously. The reason behind choosing logistic regression is its mathematical flexibility. Now logistic regression is extensively used where a relationship has to be formed among the input variables and a dichotomous output to be obtained.

In Fig. 69.1, mathematically it is represented as:

$$Y = \theta_0 + \theta_1 X \tag{69.1}$$

X is the input used as feature and theta0 and theta1 are parameters of the regression model.

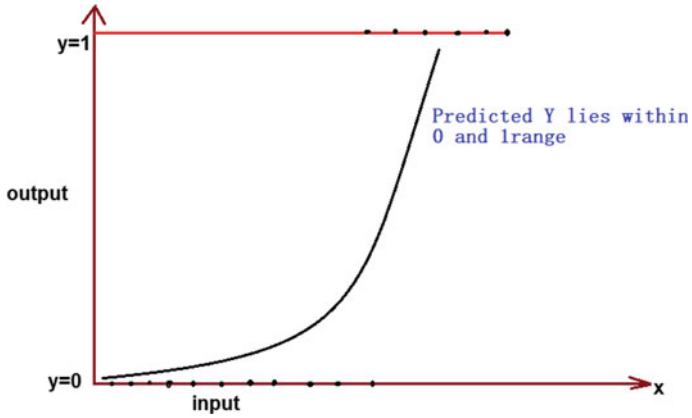


Fig. 69.1 General logistic regression curve

69.2 Literature Survey

Alshouiliy et al. [1] in this paper, by using Azure machine learning (AzureML) platform, introduced a generic IoT. It has been designed by Microsoft to ensure accuracy of the prediction. Iris Calibo et al. [2] in this paper take past nine years data to estimate the contributing features. By using linear regression as visualized on a trend line model, its objective is to focus on credit risk feature selection scoring on loan performance, used for the small and medium enterprise technology upgrading program region. This takes historical data such as debt on asset ratio (DAR), liability ratio ((LR), net profit margin ((NPM), return on investment (ROI), and year. Mashanovich et al. [3] introduced different steps for processing data with multidimensional data of borrowers which is the major step through the data mining and making different group of characteristics. The technique of testing the strength and predictive data throughout the mining view is the objective of the paper which is published at big data zone (DZone) for variable selection and big data analytics in credit score modeling. Vaidya's [4] objective in the paper is to apply machine learning tools and concepts to real-world challenges which is to develop the analytical models. Many industries are now using the capability of AI to grow and get the profit in business. Yadav et al. [5] analyze the credit risk and fraud detection which is really helpful to bank for investors who invest money or give loan to the borrower in the high time; it gives a crystal clear idea. So in Hadoop methodology, it provides different tools for analyzing data in HiveQL and storing data in HDFS.

69.3 Performance of Logistic Regression and Algorithm Used

In our model of logistic regression, we are only focused about the probabilistic value of the outcome-dependent values. Here, we use binary classifier; it will find only success (yes) or failure (no). The value of Y from Eq. (69.1) describes two things: probability of success (p) or failure ($1 - p$). Also Y must satisfy this condition to come out of logistic regression. Y must be positive and less than equal to 1. Here, we choose our threshold (the probability of success) is more than or equal to 50% (<=0.5), then we can consider the applicant who will get the loan approved. Endmost to come by a better accuracy, we will calculate the model's correctness by using "confusion matrix" which is nothing but the ratio between the actual value and the predicted value. The combination of confusion matrix is true positive (predicted true and actual true), true negative (predicted negative and actual negative), false positive (predicted positive and actual false), and false negative (predicted negative and actual false); usually, it is represented in tabular form to avoid over-fitting (noise of training data) (Fig. 69.2).

Fig. 69.3 describes the algorithm which we have used and tried to getting our simple and best-predicted accuracy model for loan approval status.

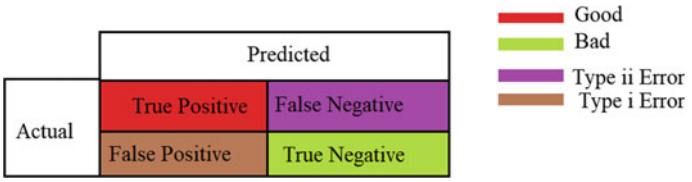


Fig. 69.2 Combination of confusion matrix

69.4 Experimental Setup and Result

In this section, we introduce our experimental setup. In the first step, we download the dataset from Kaggle, then we upload it to our work, and we need to clean and munging it to prepare for our model. There are more steps to be done on the data to be fit with our work and our model. The data contain different attributes that we are choosing attributes for prediction of loan approval to fit into the model. Table 69.1 shows all the attributes which are choosing to fit into the model of loan approval status. Basic information content of the borrowers like the gender, education, relationship status, whether or not self_employed, income status, financially dependent or not, loan amount, and credit_history is considered.

A box plot is shown in Fig. 69.4 for a statistical analysis of the input attributes; box plot or it is known as five-number summary which presents a visualization like max, min, median, lower quartile, and upper quartile point of the data. It is useful in indicating whether a distribution is skewed and whether there are any potential outliers in the dataset, and often used in explanatory data analysis (EDA). This graph is used to show the distribution, its central value, and its variability.

It is pretty tough to predict exactly which people will get the loan, so instead we calculate the probability or likelihood of a person getting it using logistic regression here. For that, we decide a threshold or cut-off value, and then, people with probability higher than that will get the loan or else not. We decided to create a plot (Fig. 69.5) with accuracy and F1 scores for both training and testing set against various threshold values to get the optimal threshold value for our algorithm. From the graph below, we found out that the optimal value for decision threshold is 0.4.

In this Fig. 69.6, it shows the confusion matrix on test data. Here, we obtained the accuracy 86%. To get the F1 measure, we have to calculate the recall (correct prediction out of all positive value) and precision (actual positive and we predict correct out of all positive value); we got 90% F1 measure.

69.5 Conclusion and Future Work

Millions if not billions of people and companies apply for loans in a day all over the world. Bank loans are one of the many parameters that help reveal the performance

Algorithm:

Step1:

input Dataset = Loan_Data

if Loan_Data=Null or Duplicate

remove variable

Step2:

Statistical Analysis of data

Mean,Std,Min,Max,25%,50%,75%

Step3:

Normalize the data
Min-Max normalization

Step4:

Split the Dataset
training_data=Loan_Data*T/100 (T is threshold value)
testing_data=Loan_Data - training_Data

Step5:

Find minimum error average rate R
R= n (n is threshold)
find the best R (minimum avg rate)

Step6:

Select model(Logistic Regression)
Predict= loan_status class_label

Step7:

Evaluating the result
Accuracy=Correct tests / Total tests

Fig. 69.3 Algorithm used in regression model

and credit risks of a financial institution. It not only helps borrowers to take money but also helps investors to invest in order to get higher interest rates. So, credit risk analysis and recommendation system focuses on the identifying the factors that would help the financial institutions to get a credit risk scoring using analytics. In today's day

Table 69.1 Contents of dataset

	Loan_ID	Gender	Married	Dependents	Education	Self_employed	Applicant income	Co-applicant income	Loan amount	Loan_amount_term	Credit_history
0	LP001002	Male	No	0	Graduate	No	5849	0.0	NaN	360.0	1.0
1	LP001003	Male	Yes	1	Graduate	No	4583	1508.0	128.0	360.0	1.0
2	LP001005	Male	Yes	0	Graduate	Yes	3000	0.0	66.0	360.0	1.0
3	LP001006	Male	Yes	0	Not graduate	No	2583	2358.0	120.0	360.0	1.0
4	LP001008	Male	No	0	Graduate	No	6000	0.0	141.0	360.0	1.0
5	LP001011	Male	Yes	2	Graduate	Yes	5417	4196.0	267.0	360.0	1.0
6	LP001012	Male	Yes	0	Not graduate	No	2333	1516.0	95.0	36.0	1.0
7	LP001014	Male	Yes	3+	Graduate	No	3036	2504.0	158.0	360.0	0.0

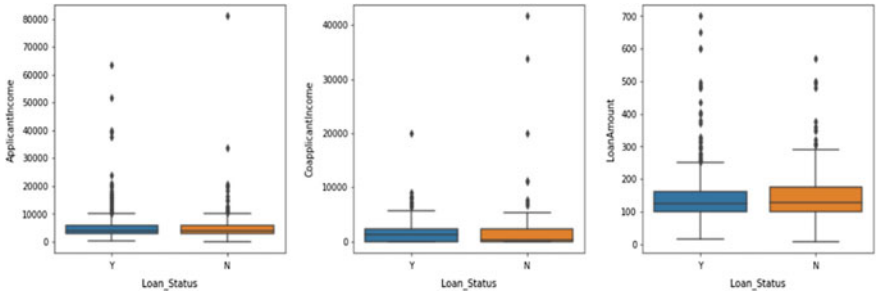


Fig. 69.4 Box plot grouped by applicant_income, co-applicant_income, and loan_amount

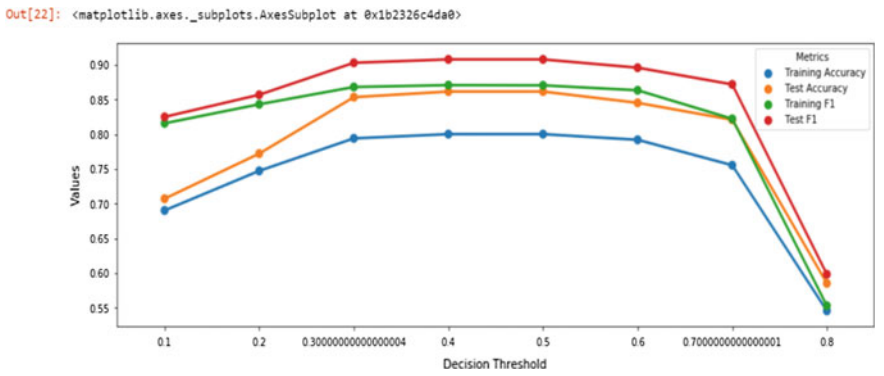


Fig. 69.5 Decision threshold for logistic regression

Fig. 69.6 Confusion matrix on test data and mathematical calculation

Out[25]:

Predicted	0	1	All
True			
0	22	16	38
1	1	84	85
All	23	100	123

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{All}} = 0.86$$

$$\text{F1 Measure} = \frac{2*\text{Recall}*\text{Precision}}{\text{Recall} + \text{Precision}} = 0.90$$

and age, both the sales representatives and software developers face increasing challenges in financial services which is complex and requires high involvement. Sales representatives are responsible for identifying tailored product or service recommendations for customers, and software developers are responsible for developing and maintaining that sales support environment. By using regression, we can easily manage data which is taking a less time to identify a particular customer information, and according to the request of customer, we can perform the best prediction model. For logistic regression, we need independent variables for assessment, and otherwise, the regression model tends to over weigh of the dependent attributes. In the future work, we will be doing the predictive analysis by using more techniques, and we will try to have a comparative analysis which will decide which model is the best.

References

1. Kanaujia, P.K.M., Pandey, M., Rautaray, S.S.: A framework for development of recommender system for financial data analysis. Int. J. Inf. Eng. Electron. Business (2017)
2. Kanaujia, P.K.M., Pandey, M., Rautaray, S.S.: Real time financial analysis using big data technologies. In: International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) (2017)
3. Alshouiliy, K., AlGhamdi, A., Agrawal, D.P.: AzureML based analysis and prediction loan borrowers creditworthy. In: 2020 3rd International Conference on Information and Computer Technology (ICICT) (2020)
4. Ferreira, L.E.B., Barddal, J.P., Gomes, H.M., Enembreck, F.: Improving credit risk prediction online peer-to-peer (P2P) lending using imbalanced learning techniques. In: 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI) (2017)
5. Bindal, A., Chaurasia, S.: Predictive risk analysis for loan repayment of credit card clients. In: 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT) (2018)
6. Arun, K., Ishan, G., Sanmeet, K.: Loan approval prediction based on machine learning approach. IOSR J. Comput. Eng. (IOSR-JCE) **18**(3), Ver. I, 79–81 (2016). e-ISSN: 2278-0661, p-ISSN: 2278-8727
7. Deng, T.: Study of the prediction of micro-loan default based on Logit model. In: 2019 International Conference on Economic Management and Model Engineering (ICEMME) (2019)
8. Leung, C.K., Kajal, A., Won, Y., Choi, J.M.C.: Big data analytics for personalized recommendation systems. In: IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress (2019)
9. Calibo, D.I., Ballera, M.A.: Variable selection for credit risk scoring on loan performance using regression analysis. In: IEEE 4th International Conference on Computer and Communication Systems (2019)
10. Natasha Mashanovich, N.: Variable selection and big data analytics in credit score modeling, DZone Big Data Zone. Available at: https://dzone.com/articles/credit-scoring-part-4-variable-selection (2017)
11. Arutjothi, G., Senthamarai, C.: Prediction of loan status in commercial bank using machine learning classifier. In: Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS 2017) IEEE Xplore Compliant—Part Number: CFP17M 19-ART, ISBN: 978-1-5386-1959-9 (2017)

Chapter 70
Efficient Load Balancing in Distributed Computing Environments with Enhanced User Priority Modeling



Maniza Hijab and Avula Damodaram

Abstract In the distributed computing environments, user-submitted job details such as the expected job completion time estimates are prone to inaccuracies. These inaccurate details force the system to under-perform due to ineffective allocation of processing resources. Addressing the concerns that crop-up due to the presence of these ill-defined user-provided parameters is an important task, which is very much required to get the jobs executed using optimal number of resources. At the same time, studying the ways of deriving accurate job runtimes benefits us in attaining improved load balancing results. The work proposed in the present study tries to apply an enhanced user priority model, taking into account the penalty and aging of job requests. It also combines least load variance method to improve the load balancing across grid resources. Simulation results obtained using realistic workloads bring forward the load balance efficiency of the studied scheme. This also confirms that to consider and infer of improved modeling of these user priority provides gains that outperform the methods void of it. The results attained are equated and contrasted with prevailing algorithms in terms of parameters like response time, wait time, tardiness, and average resource utilization in high-performance computing environments.

70.1 Introduction

High-performance computing (HPC) is unarguably the most widely used computing paradigm to solve complex real-time problems. It uses concurrently executing networked-computing resources in large numbers. It has the capability to bring to life unrelenting performance with simultaneous using thousands of networked-computing servers. Domains that use HPC include healthcare, engineering, space

M. Hijab (✉)
Muffakham Jah College of Engineering & Technology, Hyderabad, India

A. Damodaram
School of Information Technology, JNTUH, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_70

research, urban planning, finance, and business to name some [1]. The intrinsic property of HPC applications is the requirement for soaring processing speeds. This requirement for high computational speeds comes at a cost that is time varying [2].

It is in light of a legitimate concern for HPC clients to tune proficiently to utilize compute resources. Along these lines, the HPC users are centered on the fact to decrease the resource expense while getting optimal execution results.

70.2 Related Work

Various research works have addressed the concerns of modeling the techniques to compensate the inaccurate task execution times submitted by users [3–6]. The main offerings of this article can be summarized as: (i) improving on the existing method of using penalty and aging-based task execution times accuracy modeling [4] and (ii) developing a hybrid task scheduling algorithm for improved load balancing in HPC environments.

Numerous relevant associated works described the user runtime approximations of submitted jobs are frequently imprecise. For instance, as specified in [5] by Tsafrir, more than few of the studies taken up provided startling and counter-intuitive outcomes as well. Despite the fact that some scholars inferred that imprecise estimates are typically desirable over precise ones, other works showed that runtime performance is indifferent to correctness of user's runtime approximations [3, 7].

The work as reported in [1] presents that in four reports of dissimilar high-performance computers, on an average of 55% of submitted jobs ended in using below 20% of their demanded CPU time. Further characteristics were too stated, for illustration, the relationship amid unsuccessful jobs and accurateness, and between job span and correctness. The influence of user job execution time approximations has been a substance of learning in numerous reported studies. The observed study stated in [3] showed that first-come-first-served is insensitive to user runtime approximations.

Nevertheless, using precise runtime approximations definitely will enhance performance on allocated policies that give preference to short jobs, similar to SJF.

In [3], the author examined three realistic task assignments and anticipated a workload generation model in view of reservations. We calculated enhancements of around 30% in processing performance and around 15% in energy consumption.

70.3 Proposed Work

The proposed work studies the implications of using an improved user priority model based on penalty and aging of job requests combined with least load variance method to improve the load balancing across grid resources. The Penalty-based Scheduling

Table 70.1 PSP + Aging Method Priority Assignments

Tag name	Accuracy interval	Priority
C 1	(0.00, 0.05)	1
C 2	[0.05, 0.10)	10
C 3	[0.10, 0.15)	20
C 4	[0.15, 0.20)	25
C 5	[0.20, 0.30)	30
C 6	[0.30, 0.40)	35
C 7	[0.40, 0.52)	40
C 8	[0.52, 0.64)	43
C 9	[0.64, 0.78)	46
C 10	[0.78, 1.00]	49

Policy (PSP) with aging algorithm adjusts the users' job priority over time dynamically to eliminate the job starvation problem. To this end, the initial job priority is assigned according to the model used in [4].

It uses an aging mechanism to manipulate the job submission priority over time. The aim of the aging mechanism is to prevent jobs from starving; see Table 70.1.

For newly scheduled job

$$\text{priority}_{\text{new}} = \text{priority}_{\text{calc}} \tag{70.1}$$

For already scheduled job

$$\text{priority}_{\text{new}} = \text{priority}_{\text{calc}} + \text{priority}_{\text{curr}} * (\text{job\_waiting time}_{\text{curr}} / \text{job\_estimated time}_{\text{curr}}) \tag{70.2}$$

We combine the least variance-based load balancing method with the PSP + Aging scheme as follows.

Each grid resource in the distributed computing environment devotes a collection of machines, which are in turn, a combination of processing elements (PEs) when dealing out with user submitted jobs or requests. These processing elements, which are generally labeled as busy PEs, will not be allocated new jobs until they finish existing jobs. The ratio of busy PEs fraction to the overall number of non-busy PEs at a resource at any time contributes to the current load at that resource.

The average load across all the grid resources is computed using the loads at each resource. As and when a grid resource has adequate number of non-busy PEs to execute a user job (i.e., gridlet), it is examined for an additional check. This check is performed by relating the load at a selected grid resource to the average load across all the grid resources. The resource which is found to have the least average load variance is designated for allocated a task. In the following portion, we summarize the significant steps of the realized algorithm.

70.4 Simulation Setup

An extension to the GridSim [8] simulator is ALEA, which implements the features that are required by an innovative job scheduling simulator. The uniqueness of ALEA [9] simulation framework is that the basic structure for various scheduling algorithms, both queue and schedule based, is made readily available for customization. Further, it not only allows the customization of the scheduling algorithms but also has built-in provisions to accept various criteria that are to be used in modeling a distributed computing environment. These parameters are not limited only to common properties like job heterogeneity, resource heterogeneity, job inflow rates, or resource failures. Any customized parameter can also be included.

Algorithm 1: Load Balancing with Enhanced User Priority Model Scheme

Input: (a) A Tasks queue and (b) a set of grid Resources
Output: Task allocation to Resources with improved Load Balancing
Data: R: Resource set, G: Task queue, Ri: Resource instance, Gi: Task instance- and PEs: Processing Elements

```
for each task in G do/* Initialize task priority*/
  Calculate initial priority of Gi;
for each resource in R do/* Calculate average load */
  Find the total fraction of busy PEs in the grid;
Find the average load across the grid;
for each task in G do/*Start resource selection */
  Calculate current priority of Gi;/*update priority */
  if Gi has not enough priority to run then
     break;/* defer scheduling low priority task */
  else
    for each resource in R do
      if required number of PEs is available then
         Calculate the load difference with average load
             across grid;
    Select the resource with least average load variance;
    Update the total Fraction of busy-PEs for selected
        resource;
    Update the average load across the grid;
```

We now describe the organization of the various components of the proposed system which are held together in the computing environment. The anticipated system involves a set of heterogeneous grid resources (R). Each resource is in turn composed of a collection of machines (M). At the next level, the machines put together a collection of processing elements (PEs). Each resource has a variable number of machines (M) accompanying it. The collection of processing elements (PEs) that are run by a machine (M) have varying architectures and speeds.

The ALEA framework accompanies realistic datasets from the Gaia cluster which recorded the log whose details are elaborated in [10]. These are used to test the functionality of the implemented algorithms. One concrete example that we can infer here is the resource with name Gaia [1–60]. It means that it is has 60 machines, each one of it composed of 12 heterogeneous processing elements like the Intel Xeon

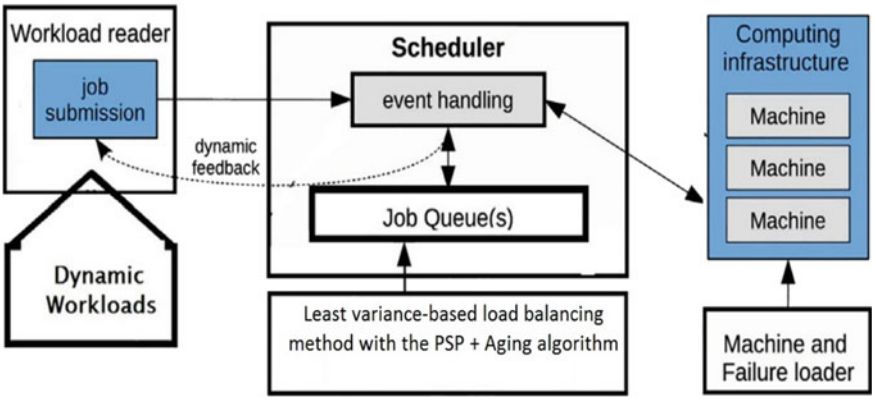


Fig. 70.1 Architecture of ALEA depicting the customized module

L5640 cores and varying speeds in the range of 2 GHz. Tasks, which are otherwise called gridlets, are scheduled on to these grid resources as and when loaded by user applications for processing. The gridlets are allocated on to the resources by means of a unified scheduler [9].

We investigated the schemes discoursed in the previous section using the said realistic data records of tasks offered in the standard workload format (SWF) that were executed on the Gaia cluster [10]. The carefully chosen real-time task logs contain the data for a period of more than a month. It has recorded a count as far as 50 thousand plus jobs allocated on to the grid resources [11].

Below given is Fig. 70.1, which depicts the important modules that constitute the ALEA simulator. The scheduling algorithms module was customized to incorporate the proposed least variance-based load balancing method with the PSP + Aging scheme as follows.

70.5 Results and Discussions

In the figure (Fig. 70.2), after the conduct of experiments with the said data records, a response time comparison is drawn for EDF, P2SDLB [11], and the ALVA (with improved user priority model) algorithm. It brings out that ALVA scheme is 31.8% more efficient than the traditional schemes.

In the figure (Fig. 70.3), with the said data records being used in the experiments, a wait time comparison is drawn for EDF, P2SDLB [11], and the ALVA (with improved user priority model) algorithm. It brings out that ALVA scheme is 89.3% better performing and efficient than the traditional schemes.

In the figure (Fig. 70.4), with the experiments revealing the insights with the said data records, a job tardiness comparison is drawn for EDF, P2SDLB [11], and

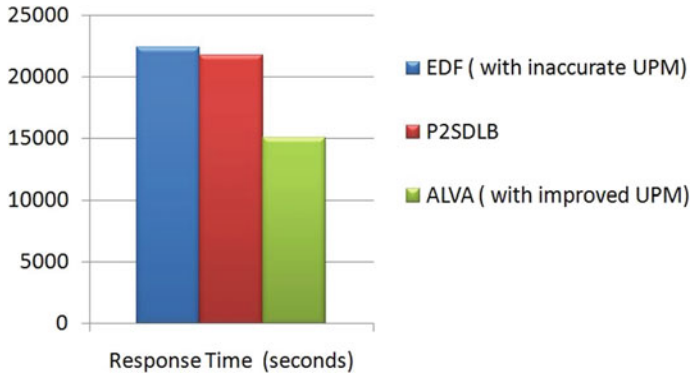


Fig. 70.2 Comparison of response time for the implemented algorithms

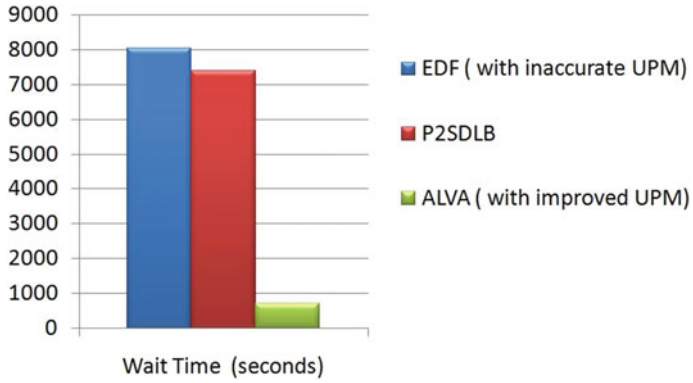


Fig. 70.3 Comparison of job wait time for the implemented algorithms

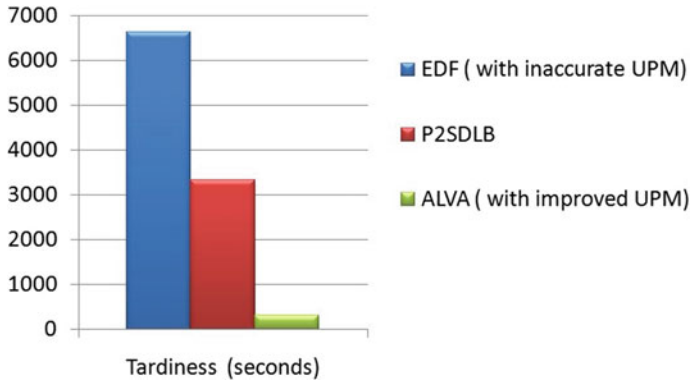


Fig. 70.4 Comparison of job tardiness for the implemented algorithms

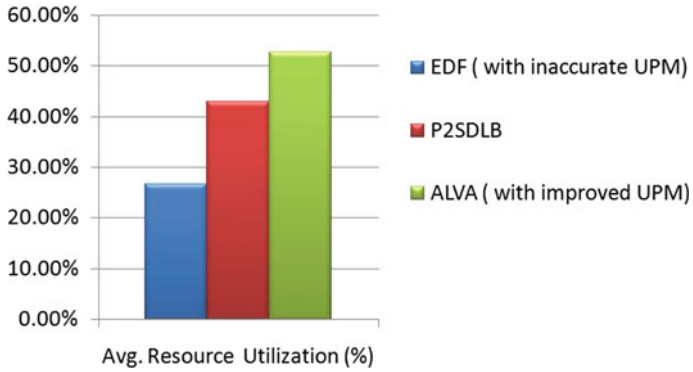


Fig. 70.5 Comparison of average resource utilization for the implemented algorithms

the ALVA (with improved user priority model) algorithm. It brings out that ALVA scheme is 91.1% proficient than the traditional schemes.

In the figure (Fig. 70.5), with the experiments revealing the insights with the said data records, an average resource utilization comparison is drawn for EDF, P2SDLB [11], and the ALVA (with improved user priority model) algorithm. It brings out that ALVA scheme is 32.69% proficient than the traditional schemes.

70.6 Conclusion

In the proposed work, to optimize the response time, wait time, tardiness, and average resource utilization, we use a least average load variance-based algorithm combined with a better user priority model. This algorithm is successful in providing a lower bound on the load across grid resources while ensuring a better response time compared to other methods proposed in literature. With the experiments revealing the insights with the said data records, the ALVA (with improved user priority model) algorithm outclasses traditional solutions in terms of improvement of resource utilization with an optimal response time, wait time, and tardiness.

References

1. Rice, M.: 14 High Performance Computing Applications & Examples. https://builtin.com/hardware/high-performance-computing-applications/. Accessed 19 June 2020
2. Balis, B., et al.: Porting HPC applications to the cloud: a multi-frontal solver case study. J. Comput. Sci. **18**, 106–116 (2017)
3. Rocchetti, N., Iturriaga, S., Nesmachnow, S.: Including accurate user estimates in HPC schedulers: ban empirical analysis. In: XXI CongresoArgentino de Ciencias de la Computación (Junín, 2015) 2015

4. Rocchetti, N., et al.: Penalty scheduling policy applying user estimates and aging for supercomputing centers. In: Latin American High-Performance Computing Conference. Springer, Cham (2016)
5. Emeras, J., et al.: Evalix: classification and prediction of job resource consumption on HPC platforms. In: Job Scheduling Strategies for Parallel Processing. Springer, Cham (2015)
6. Galleguillos, C., et al.: Constraint programming-based job dispatching for modern HPC applications. In: International Conference on Principles and Practice of Constraint Programming. Springer, Cham (2019)
7. Zotkin, D., Keleher, P.J.: Job-length estimation and performance in backfilling schedulers. In: 8th International Symposium on High Performance Distributed Computing, pp. 236–243 (1999)
8. Buyya, R., Murshed, M.: Gridsim: A toolkit for the modeling and simulation of distributedresource management and scheduling for grid computing. Concurr. Comput. Pract. Exp. **14**(13–15), 1175–1220 (2002)
9. Klusáček, D., Tóth, Š., Podolníková, G.: Complex job scheduling simulations with Alea 4. In: Proceedings of the 9th EAI International Conference on Simulation Tools and Techniques (2016)
10. ULHPC: The gaia cluster hpc at university of luxemburg. https://hpc.uni.lu/systems/gaia/. Accessed 26 March 2020
11. Thakor, D., Patel, B.: P2s_dlb: pluggable to scheduler dynamic load balancing algorithm for distributed computing environment. In: Emerging Trends in Expert Applications and Security, pp. 347–355. Springer, Singapore (2019)
12. Guim, F., Corbaĺan, J., Labarta, J.: Prediction f-based models for evaluating backfilling scheduling policies. In: 8th IEEE International Conference on Parallel and Distributed Computing, Applications & Technologies, pp. 9–17 (2007)

Chapter 71
Deep Neural Networks Based Object Detection for Road Safety Using YOLO-V3



Jalaja Tattari, Vineeth Reddy Donthi, Dheeraj Mukirala, and S. Komal Kour

Abstract Traffic safety plays a pivotal role in the society, and hence, there is a grave need for streamlining the protocols and measures taken in our country. This paper gives a comprehensive solution to the situation at hand, a new approach to 'road safety' based on seatbelt and helmet object detection using deep neural networks. The pipeline consists of using the deep learning frameworks and 'You Only Look Once V3 (YOLO-V3)', which is a convolutional neural network-based algorithm, for a hassle-free "real-time object detection." Detection systems in the past re-engineer classifiers or localizers to perform detection. The paper focuses on a single neural network applied to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. Popular image data sources such as the UCI Machine Learning Repository and Google Open Image Datasets were explored, and a dataset was created, which underwent preprocessing. The developed model is very versatile, compact, and was tested on various edge cases after which the model achieved an overall accuracy of 86.3% which was on a higher note compared to the previous works. The program can be readily deployed on CCTV, Webcams, or local computers.

71.1 Introduction

The cutting-edge technology achievable using the deep learning (DL) domain has been giving the mankind glorious results in the recent past. It is in this context that one should think about the possible applications of deep learning with special emphasis on computer vision. The paper revolves around using efficient computer vision algorithms for object detection for the betterment of the current road safety scenario in the country. It was reported by the road transportation authority that over 5 lakh traffic-related accidents occur every year, out of which 1 lakh are deaths. It is a very well-known fact that not wearing helmets and seat belts while driving on the road is a punishable offense under the law, and the offenders are penalized;

J. Tattari (✉) · V. R. Donthi · D. Mukirala · S. Komal Kour
Vasavi College of Engineering, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_71

however, there are many offenders out on the roads, still very reluctant to follow the safety protocols. Governments have been doing a spectacular job raising awareness about the guidelines and also establishing violation detection systems that either use manual detection or algorithms that are less accurate and are not fast.

Technology was static and stagnant earlier, but it is now aggressively dynamic. We are headed toward a future in which machines take over humans with decision-making abilities, hence making models with high output capabilities is of paramount importance. The road safety and transportation domain needs a massive restoration in the pillars of current technologies and techniques. The paper aims at building a real-time, reliable, and a relinquishing model using the "You Only Look Once" (YOLO-V3) algorithm for helmet and seatbelt detection in two wheelers and four wheelers, respectively.

71.2 Related Work

In the paper by Babu et al. [1], they have accomplished object detection using machine learning. The training was carried out with different algorithms like support vector machine (SVM), gradient-boosted trees, etc. The primary aim of the model is to detect the presence of a motorcycle and if there is one, to detect the presence of a helmet on the rider.

In the paper by Nguyen et al. [2], small object detections were implemented using different deep learning algorithms and have proved that YOLO-V3 has faster and accurate predictions. The model was trained on the common objects in context (COCO) dataset.

In the paper by Rohit et al. [3], the videos from a traffic surveillance camera are fed into an object detection model. When the model detects the human class, the output is given to an image classifier which will compare reference image with the new output image and tells if the rider and co-passenger are wearing a helmet or not based on a classification model.

In the paper by Huang et al. [4], they used the YOLO-V3 algorithm to perform traffic flow detections on surveillance videos. The model aims at detecting moving vehicles for improvement of transportation facilities. A dataset containing a wide variety of weather conditions and types of vehicles was used to ensure accurate predictions.

71.3 Proposed Work

A deep learning-based method has been set forth to curb today's alarming increase in road accidents, especially in India. The YOLO-V3 algorithm was used to ensure better accuracy and speed (Fig. 71.1).

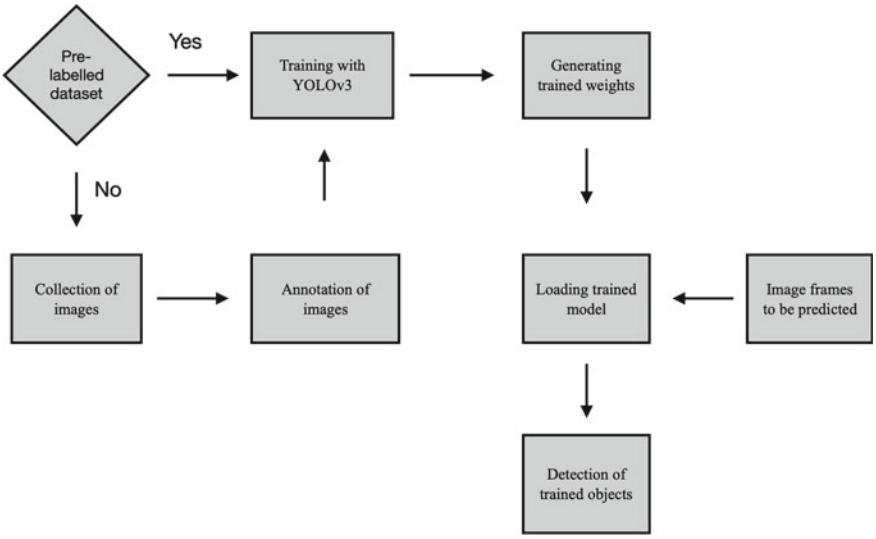


Fig. 71.1 Proposed work flowchart

71.3.1 YOLO-v3

There are many object detection algorithms, YOLO being one of the fastest. Out of all the algorithms, this is opted often because it maintains higher accuracy for real-time detection as well. It detects multiple objects in the feed unlike the other image classification algorithms which only classify the image as a specific class. It finds the objects in an image and draws bounding boxes around them. Exact boundaries of objects can also be drawn using instance segmentation.

The ability of YOLO-V3 to apply a single neural network to the full image rather than apply a model to an image at multiple locations and scales makes it stand out from the other algorithms. This unique ability makes it 1000× faster than R-CNN and 100× faster than Fast R-CNN.

This method divides the input image into small-squared grid cells. A grid cell is held responsible for detecting an object if the center of an object falls in it. Each grid cell predicts the position information of bounding boxes and computes the scores corresponding to the boxes [5].

The feature extractor used in YOLO-V3, i.e., Darknet-53, consists of 53 convolutional layers, while the remaining levels are called resident layers [6].

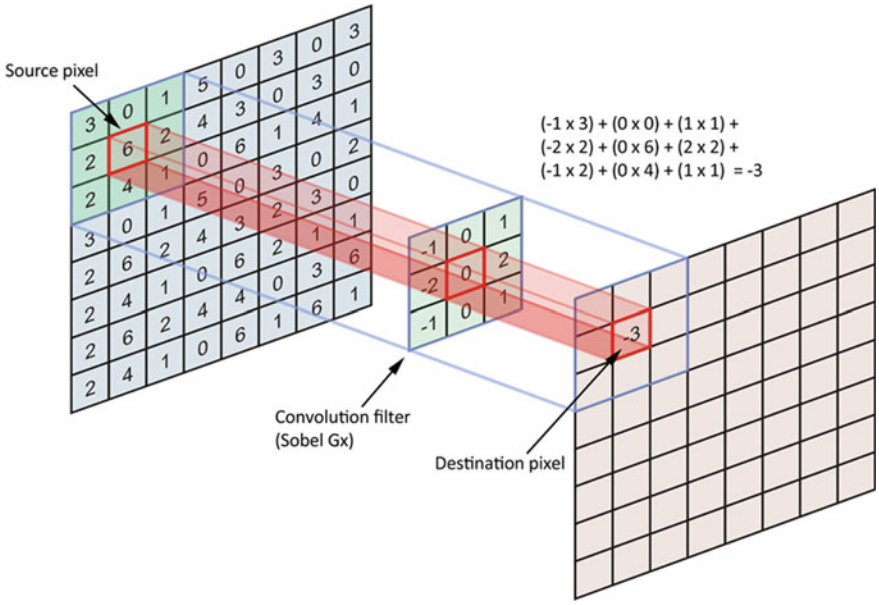


Fig. 71.2 Convolution filter 3D representation

71.3.2 CNN

"A convolutional neural network (CNN) is a deep learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate one from the other" [7].

The employment of convolutional neural networks gives us the ability to automatically learn a large number of filters in parallel specific to a training dataset under the constraints of a specific predictive modeling problem, such as image classification. This ability helps us detect highly specific features anywhere on input images [8]. It converts an image into matrices of numbers. A kernel/filter is a 3×3 matrix which is convolved over the entire matrix as shown in Fig. 71.2.[1]

ReLU is a nonlinear activation function like tanh or sigmoid. It takes matrix from a reduced image as input and replaces all the negative values with zero keeping the other values constant.

The softmax activation is normally applied to the very last layer in a neural net. The importance of softmax lies in converting the output of the last layer in a neural network into a probability distribution.

[1] Source: https://datascience.stackexchange.com/questions/23183/why-convolutions-always-use-odd-numbers-as-filter-size.

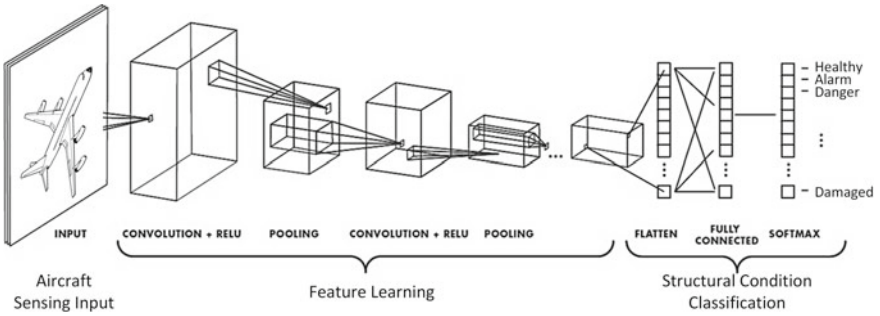


Fig. 71.3 CNN workflow

The entire functioning of CNN's is summarized as shown in Fig. 71.3 [9] [2].

71.3.3 Input

A dataset consisting of different types of helmets and seatbelt altogether 5000 images were taken as input. COCO dataset was used as it is an excellent object detection dataset with 80 classes, 80,000 training images, and 40,000 validation images. A pre-labeled dataset was used to avoid blunders while annotating images, i.e., labeling the images with their respective classes. The dataset was divided into training set and test set in the ratio of 4:1, i.e., for every four images trained, the model will be tested on one image to prevent overfitting of the model.

71.3.4 Training

Google Colab was used for training the model to make use of Tesla K80 GPU (Graphics Processing Units). The model was trained for 3000 iterations to clinch greater accuracy. Though the dataset plays a major role in achieving higher accuracy, preparing configuration files is also equally important. In the data file, the classes must be given the number of classes that you are training the model on, in this case two. Using a piece of code, text files are generated for the training and test sets; these text files must be stored in the variables train and valid in the.data file. This file also must contain another variable "names" which must have the.names file.

[2]Tabian, Iuliana; Fu, Hailing; Sharif Khodaei, Zahra. 2019. "A Convolutional Neural Network for Impact Detection and Characterization of Complex Composite Structures." *Sensors* 19, no. 22: 4933.

The .cfg files must be chosen based on the GPU available. Since Google Colab provides GPU, yolo.cfg file was used. The GPU must be set to 1 in the runtime type to fully utilize the hardware. In this file, the following parameters were changed according to our model: batch, subdivisions, filters, and classes.

The filters used in the convolutional layer were calculated using the formula,

$$\text{filters} = (\text{number of classes} + 5) * 3 \tag{71.1}$$

making it 21 in this model. It took 8–9 h approximately to train the network with the above parameters. To save the weights, more often the settings were changed accordingly. The weights generated after 3000 iterations were used to detect the trained objects in random traffic images. [10]

71.4 Experiments and Results

71.4.1 Experiments

The training was carried out on a custom stitched dataset which was examined carefully for imbalance. The dataset consisted around 5000 images which was the backbone of the training process. We have maintained the batch size to be 64 even though batch size 32 would have made our model converge even faster. A backup folder was created in the parent directory to save weights to this folder automatically after every 200 iterations until 1000 iterations and for every 1000 iterations thereafter. We finished training in about 3 days and stopped the training at about 3000 iterations because the model converged to a cogent mAP accuracy of 86.3%.

71.4.2 Results

A few examples of the output object detector are shown in Figs. 71.4 and 71.5.

71.5 Conclusion and Future Work

From the above result and analysis, it is very evident that the generated model is very adaptive in nature. Computer vision algorithms are the need of the hour in the road safety domain in our country. The proposed model is dressed to the nine and has reached higher accuracies than the existing frameworks in the realm of object detection. The role of hyperparameter tuning and optimization is very essential for clinching higher performances. This paper is capable of paving a tenacious path in



Fig. 71.4 Helmet detection



Fig. 71.5 Seatbelt detection

the future of deep learning in several areas like self-driving vehicles, monitoring crops for disease detection, medical diagnostics, etc., one of the major applications being penalizing traffic offenders to ensure public safety.

References

1. Ramesh Babu, D.R., Rathee, A., Kalita, K., Deo, M.S.: Helmet detection on two-wheeler riders using machine learning. In: ARSSS International Conference (2018)
2. Nguyen, N.-D., Do, T., Ngo, T.D., Le, D.-D.: An Evaluation of Deep Learning Methods for Small Object Detection, Hindawi (2020)
3. Rohith, C.A., Nair, S.A., Nair, P.S., Alphonsa, S.: An efficient helmet detection for MVD using deep learning. In: Proceedings of the Third International Conference on Trends in Electronics and Informatics (2019)
4. Huang, Y.-Q., Zheng, J.-C., Sun, S.-D., Yang, C.-F., Liu, J.: Optimized YOLOv3 algorithm and its application in traffic flow detections. Appl. Sci. (2020)
5. Zhao, L., Li, S.: Object detection algorithm based on improved YOLOv3, Computer Science and Engineering (2020)
6. Vidyavani, A., Dheeraj, K., Rama Mohan Reddy, M., Naveen Kumar, K.H.: Object detection method based on YOLOv3 using deep learning networks. IJITEE (2019)
7. Saha, S.: A Comprehensive Guide to Convolutional Neural Networks—The ELI5 way, TowardsDataScience (2018)
8. Brownlee, J.: How do convolutional layers work in deep learning neural networks? Deep Learn. Comput. Vision (2019)
9. Tabian, I., Fu, H., Sharif Khodaei, Z.: A convolutional neural network for impact detection and characterization of complex composite structures. Sensors **19**(22), 4933 (2019)
10. Masurekar, O., Jadhav, O., Kulkarni, P., Patil, S.: Real time object detection using YOLOv3. IRJET (2020)

Chapter 72
Log-Based Authentication for Cloud Environments



Nalini Subramanian and G. Shobana

Abstract In the world of cloud era, several organizations achieve benefits in terms of capital expenditure and operational expenditure. Security in adopting cloud is one of the important challenges that need to be addressed for several organizations. In order to increase the security, cloud should provide an advanced authentication framework for accessing its resources. This paper introduces the need for advanced authentication, based on the security issues encountered in cloud computing at various levels, since advanced authentication provides the logging functionality. When a local account is used to log on to a computer, that computer performs both logon and authentication. When a domain account is used to log on to a computer, the computer being accessed makes the connection, but a domain controller in the domain that owns the user account authenticates. So, this paper explores log-based advanced authentication, and it also depicts how the usage of dimensionality reduction technique and support vector machine (SVM) reduces the complexity of an authentication framework without sacrificing the accuracy.

72.1 Introduction

Data is at the heart of the crypto-cloud, and data breach is identified as the number one threat by the cloud security alliance (CSA). The life cycle of data from source to destruction must be secure, as storing data in a remote location and the concept of multi-tenancy can lead to data leaks. Data security is classified into data in transmission and data at rest. Inactive data attracts hackers more than transmitting data. Data security can be broadly classified into three basic levels, namely the communication level, the processing level, and the service level agreement (SLA) level.

N. Subramanian (✉)
Prathyusha Engineering College, Tiruvallur, India

G. Shobana
Department of Computer Applications, Madras Christian College (Affiliated to University of Madras), Chennai, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_72

Typically, authorized users can access the resource from the server once they are authenticated by the server. The goal is to provide a strong authentication framework for cloud environments. The authenticators used and their combination play a very important role. When a unique authenticator is used, it can easily be forged. This leads to a single point of failure. A combination of these is used because a one-factor authentication system often fails to provide high security. Multi-factor authentication is required to improve security depending on the type of application. Simply increasing the number of authentication levels alone does not increase security. On the contrary, it involves costs, time, space, etc., considerable to provide strong authentication. Three fundamental elements must be combined: what the user knows, what he has, and what he is. This multi-layered authentication framework enables secure and strong authentication for cloud environments.

Log details kept on the server are used to prevent system performance degradation and to monitor critical events. Logs are local files saved when an event occurs in the system. Events can be the insertion of a file, the deletion of a file, and its access or its modification. Security events are data or system security-related events, and the corresponding log file is called the security log. Event logs can be classified as application log, system log, and security logs, among which this work focuses on the security log, which is related to system security. An invalid connection or disconnection falls within these criteria.

The number of failed login attempts prevents or reduces data theft, giving it the highest priority for authentication. Logging of all events is not recommended as it exposes the data to hackers. Depending on the type of application, only critical events require regular monitoring. The user details provided as input along with specific information such as username, date and time, IP address, status, URL, and number of login attempts are used to improve security and authentication. As the total number of parameters increases, a dimensionality reduction technique such as PCA is required to reduce system complexity. Otherwise, it leads to false negatives.

72.2 Related Work

In the article [1], the authors presented an approach called log-based anomaly detection (LAD). The proposed approach finds the communication rules used in mobile ad-hoc networks using the logs of router. Reasoning rules are used to detect both active and passive attacks triggered during the phase of routing. Using Merkle hash tree, the files are secured, and confidentiality of the nodes is improved. The methodology detects malicious nodes without the usage of special nodes nominated for verification. This approach is implemented in NS3 using router environment to prove that mobile ad-hoc security is increased. It is proved that the accuracy rate is increased from 2.28 to 29.22%. From the results, it is confirmed that the proposed approach consumes little memory and time.

Allani and Ghannouchi [2], authors proposed an approach to evaluate the business process models as well as analysis of event logs based on the previous execution.

The models evaluated are written in BPMN2.0 (Business Process Model Notation) and executed using activity explorer online service. In order to verify the results, the experiments have been done on two different business processes.

The process of implementing the PKI improvements assured on logs has been proposed by [3] in two ways. First, the PKI influence model should encourage each other to implement a PKI enhancement and determine that potential PKI enhancements should focus their initial efforts on the browser vendor implementation. Two things, as a promising vendor-based solution, the authors implemented status filters, here use a Bloom filter to monitor the status of the distribution and effectively defend against demotion attacks. Protocols have been improved for the current PKI TLS. The findings have led to promising implementation strategies for the log-based improvements in PKI and raise further questions for further fruitful research.

A log-based analysis tool to evaluate the computer system of Web applications was proposed by [4]. This tool integrates software registry with an infrastructure. Even though software and infrastructure issues are complicated, software engineers can overcome only system approach errors. The system approach comprises of five phases: software and infrastructure preparation, log collection, log review data, and log data discovery. To a simple local area network of a small-scale Web application system, the tool was applied. The usefulness of the proposed tool was confirmed, when a software engineer detects file system crash errors such as “404” and errors of type “no response." The tool was also partially applied to a large-scale computer system with many Web applications and a large network. Using the proofreading and tracing tool, the causes of a true authentication failure were found. The causes enlisted both infrastructure problem with a software problem which are combined. It is confirmed that errors are not caused only by an issue in software. Software engineers are only responsible to distinguish a software problem when compared to an infrastructure problem using the proposed tool.

In [5], we presented real-world examples that would lead to broken instances of the idealized AuthA protocol and the One-Encryption-Key-Exchange (OEKE) protocol. The results show that the AuthA protocol can be instantiated in an insecure manner and does not exist to distinguish between safe and insecure instances. Therefore, without a strict metric for ideal numbers, the value of demonstrable security in the ideal encryption model is limited.

The log-based intrusion detection system uses user logs from Web servers, where login information is recorded on every request. The downside to log-based Web intrusion detection is that it can only be performed after transactions have been completed; therefore, prevention of attacks is impossible [6]. Focused on user behavior, to discover the difference between a common user and a malicious user and to try to prevent the attack before it happens.

72.3 Proposed Framework

When the user logs in, they are authenticated by their name and password. At the second level of authentication, the fingerprint is used. Fingerprints can be forged using jelly wax. Therefore, the network time protocol (NTP) timestamp is used for authentication. An authentication system is strong when a combination of credentials, such as a password, a biometric key, and a session key (token) is used.

To increase the accuracy of the strong authentication framework (SAF), the log details kept on the server are used with the parameters provided by the user to provide high security as in Fig. 72.1. In addition to client details, log details can be accessed to provide an advanced authentication infrastructure. Log details can be viewed according to application requirements. Typically, they will be queued at regular intervals and used only when needed to resolve a problem. The details can be viewed when needed, in order to increase the security of the application. Using this information increases the security of the cloud server that requires authentication. The information requested can be collected from the various logs available.

There are different types of logs such as audit log, secure log, and access log. The details kept in the log file vary depending on the needs of the application. Starting with the single autonomous system and moving to a network of distributed systems, logs are kept to track and identify attackers. Any user movement, such as the file they are accessing, URL details, system usage date and time, system IP address, action/status, for example, connect/failed if parameter action or code is used to a state like 404, 200, etc., is stored. This list of parameters varies depending on the application requirements. Some secure log settings are identified as containing specific user details required for authentication. The number and type of parameters

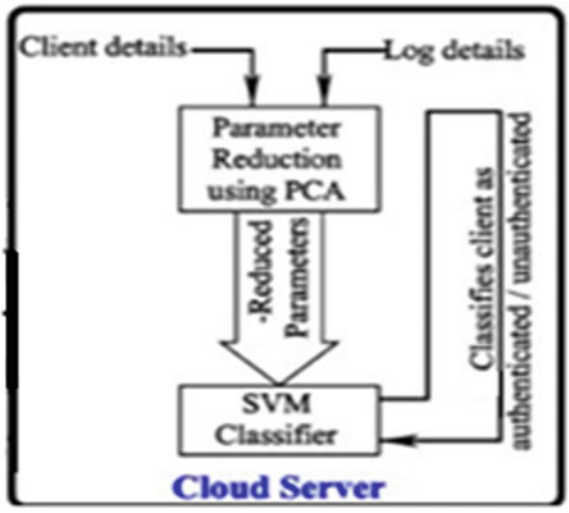


Fig. 72.1 Log-based authentication framework

vary dynamically depending on the type of user login. Variations and increases in the number of parameters lead to the need for a technique to help reduce them.

PCA is used to decrease the number of features prior to the classifier, which classifies the users as authenticated/unauthenticated.

72.3.1 Support Vector Machine

Support vector machine is a supervised machine learning model. When machine learning is implemented on a cloud server with suitable feature compression, it is certainly possible to obtain a very accurate model for predicting authenticated users. In general, the number of parameters taken should be larger to ensure accurate results. But this increases the complexity of the model along with the increase in false negatives and sometimes even false positives. This framework reduces the overall complexity and enhances the cloud environment, to permit only authenticated (true negatives) users to access resources when compared to unauthenticated (true positives) users.

The following are the overview of implementing SVM:

Step 1: Dataset to be loaded
Step 2: The loaded dataset is summarized
Step 3: It is visualized
Step 4: The model is evaluated, and predictions are made.

According to [7], the occurrence of an attack is represented by 1, and the non-occurrence of an attack is represented by 0. When attack and alarm are 0, it is identified as true negative. With attack as 0 and alarm as 1, it is false positive. If both are 1, then it is true positive, and if only attack is 1 and alarm is 0, it is identified as false negative.

72.4 Classification Phase Algorithm

The classifier algorithm is stated as follows:

Input: web_log data set

Output: Classified as authenticated/Unauthenticated users

1. Using pandas, numpy, matplotlib, and sklearn packages, SVM classification phase is implemented.
2. Predefined libraries like classificiation_report and confusion_matrix are imported.
3. Import the preprocessed dataset in.csv format (https://www.kaggle.com/shawon10/web_log_dataset#weblog.csv)

4. Dividing the data into attributes and labels. X = weblog.drop('action', axis = 1) and y = weblog['action']
5. Dividing training and testing sets, X_train, X_test, y_train, y_test = train_test_split (X, y, test_size = 0.20).
6. Model is created with SVC () of sklearn.svm, then fit () is used to train the model. Kernel parameter is set to linear, and svclassifier.fit (X_train, y_train) is invoked.
7. Predictions are performed using predict () of svclassifier as, y_pred = svclassifier.predict (X_test)
8. Model is evaluated by computing the confusion matrix and classification report, using y_test and y_pred.

Among the available models, SVM yields the highest percentage of accuracy for small datasets. The model is executed on the test data, and the results are summarized in a final accuracy score, as a confusion matrix and a classification report.

72.5 Experimental Evaluation

Once the confusion matrix is obtained, the precision, recall, accuracy, and error can be computed as in Eqs. (72.1)–(72.4). They are the parameters for the evaluation of the classifier.

$$\text{precision} = \text{true} - \text{positive}/(\text{true} - \text{positive} + \text{false} - \text{positive}) \tag{72.1}$$

$$\text{recall} = \text{true} - \text{positive}/(\text{true} - \text{positive} + \text{false} - \text{negative}) \tag{72.2}$$

$$\text{accuracy} = \text{true} - \text{positive} + \text{true} - \text{negative} / (\text{true} - \text{negative} + \text{true} - \text{positive} + \text{false} - \text{negative} + \text{false} - \text{positive}). \tag{72.3}$$

$$\text{error} = 1 - \text{accuracy}. \tag{72.4}$$

The confusion matrix and classification report for the sample of 100 records are tested with the model. Accuracy, precision, and recall are computed for 100 users. The number of false negatives is reduced when the number of record size increases as in Table 72.1. This is done from the two classes of input, namely authenticated and unauthenticated. TN = 13, FP = 4, FN = 5, and TP = 18 form the list of data, the model interpreted accurately, and accuracy is 77.5%. The precision will be 72%, and recall is 76%

Increase in accuracy is seen with increase in the number of users. Accuracy is 0.77 initially for a record size of 100, as the record size increases the accuracy also

Table 72.1 Computation of the parameters

Records	Precision	Recall	Accuracy
100	0.72	0.76	0.775
200	0.89	0.75	0.81
300	0.89	0.8	0.85
400	0.95	0.84	0.9
500	0.97	1	0.96

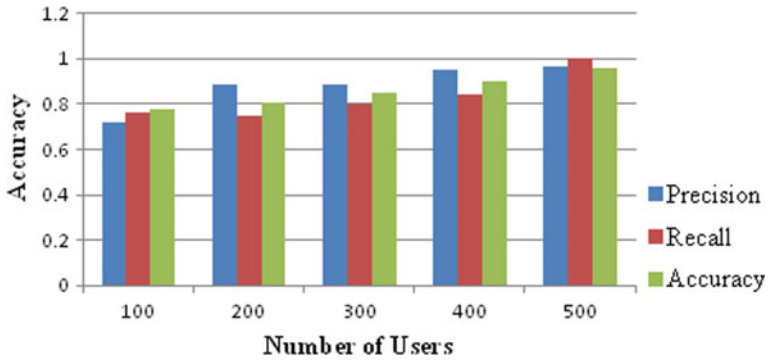


Fig. 72.2 Precision, recall, and accuracy

increases. It is 0.96 which is 96% when the record size is 500, precision is 0.97, and recall is 1 as shown in Fig. 72.2.

72.6 Conclusion

Log security and reliability are important for effective system maintenance and monitoring. This leads to the conclusion that logs are an important source for the security control system, as hackers primarily modify log data to hide their invasive behavior. The log file includes the person's personal information, so it should be protected. The security and authenticity of the log are guaranteed to protect against attackers. With the help of using dimensionality reduction technique, large numbers of redundant and irrelevant features are reduced. 'Support vector machines' work on smaller datasets, compared to other models. They can be stronger and more powerful as building models. This proposed work accurately classifies users as authenticated and unauthenticated. It includes SVM for classification and PCA for feature selection to improve accuracy and reduces the complexity. The false positive rates have been eliminated when compared to other models. The security has increased, once unauthenticated users are limited by a framework.

References

1. Li, T., Ma, J., Pei, Q., Shen, Y., Sun, C.: Log-based anomalies detection of MANETs routing with reasoning and verification. In: Proceedings, APSIPA Annual Summit and Conference, pp. 240–246 (2018)
2. Allani, O., Ghannouchi, S.A.: Verification of BPMN 2.0 process models: an event log-based approach. In: Conference on ENTERprise Information Systems/International conference on Project MANagement/Conference on Health and Social Care Information Systems and Technologies,CENTERIS/ProjMAN/HCist 2016; Proc. Comput. Sci. **100**, pp. 1064–1070 (2016)
3. Matsumoto, S., Szalachowski, P., Perrig, A.: Deployment challenges in log-based PKI enhancements, pp. 1–7 (2015)
4. Hanakawa, N., Obana, M.: A log-based trace and replay tool integrating software and infrastructure. Int. J. Softw. Eng. Appl. (IJSEA) **9**(4), 1–19 (2018)
5. Zhao, Z., Dong, Z., Wang, Y.: Security analysis of a password-based authentication protocol proposed to IEEE 1363. Theoret. Comput. Sci. **352**, 280–287 (2006)
6. Chanthini, S., Latha, K.: Log based internal intrusion detection for web applications. Int. J. Adv. Res. Ideas Innov. Technol. **5**(3), 350–353 (2019)
7. Subramanian, N., Andrews, J.: Strong authentication framework using statistical approach for cloud environments. In: Concurrency and computation: Practice and Experience (2018). ISSN: 1532-0634. https://doi.org/10.1002/cpe.4870

Chapter 73
Development of a Wireless Sensor Node for Electrical Power Distribution Monitoring System



Sanjay Chaturvedi, Arun Parakh, and H. K. Verma

Abstract Today, systems around us are expanding rapidly every day, whether it be communication systems, power distribution systems or security systems, or any other interwoven system in our daily lives. With the increasing size, these systems are becoming complex to manage as they are interdependent on several parameters. This problem of managing systems is handled very well by incorporating smartness into these systems. A system can be made smart by using current technologies like IoT and WSN. Where in IoT, the sensors are Internet-enabled and send data directly over the Internet, and the case with WSN is different in the sense that WSN does not need to have any connection to the Internet and can be used independently. The use of WSN-based systems requires wireless sensor nodes to sense the parameters of the system under observation. The available works present different ways of developing the sensor nodes. Still, the common problem with them is the cost of development and the complexity and steep learning curve to reproduce those sensor nodes within the academia for research purposes. In this work, we present a simple approach to design a wireless sensor node for application in a power monitoring system. We built the sensor node around an open-source development board, Arduino Uno, using simple circuits and readily available sensors and used ZigBee technology, an open standard, for wireless communication. We implemented the developed sensor node in our work 'wireless sensor network-based power monitoring system'. Here, we measure three critical parameters at a power distribution point viz voltage, current, and frequency. The measured values are transmitted to the central server from where the distribution system is monitored. The operating voltage of power lines varies between 230 and 245 V, and the current varies to a maximum of 30 A in our case. The frequency ranges from 49.8 to 50.25 Hz. The results obtained in implementation have a marginal error with an approximate mean square error of less than 0.5% of the parameter's maximum operating value. These results have validated our developed sensor node and the power monitoring system.

S. Chaturvedi (✉) · A. Parakh · H. K. Verma
Shri G. S. Institute of Technology & Science, Indore, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_73

73.1 Introduction

A wireless sensor network is a network of many sensing nodes, which are spatially distributed. These nodes continuously monitor the system's parameters under observation and relay the measured data to a central server over a wireless communication channel. These sensor networks are used in the monitoring of health [1], monitoring of power grids [2], detecting fire [3], in energy management systems [4], and for factory automation [5], to name few.

A wireless sensor network's three main components include sensor nodes, wireless communication link, and a central unit. The sensor node is based on some processing element like a microcontroller, which, when integrated with sensors and a wireless transceiver, forms the wireless sensor node. Wireless communication is done either by the use of available wireless network services or using an ad-hoc network. And the central unit is generally a computer responsible for coordinating with all the nodes in the sensor network and managing the data received from the sensor nodes [6]. This paper presents the development work of one of the main components of a wireless sensor network, the wireless sensor node. We find a fair amount of work in the literature where wireless sensor nodes have been built using different development boards and other approaches. Few works with different methods include the work of Prasad et al. [7], which presents a power monitoring system for smart grids. In this work, they used a power sensing IC ADE7757 to develop their sensor node. Suryadevara et al. [8] in their work have developed a monitoring and control system for household appliances. They used the processor available in XBee radio modules for the development of sensor nodes. The work done by Arora et al. [9] has presented two systems for measuring voltage, one based on PIC microcontroller PIC16F877A and the other based on the AT89S52 microcontroller of the 8051 families. This work also validates the use of 8-bit microcontrollers for AC signal sensing applications. Ahamed et al. [10] have made an energy meter using a dsPIC33F microcontroller and an external ADC. Several other works have used Rasberry Pi for developing sensor nodes.

The problem in the available works was the absence of a smooth transition from the theory being taught to the hardware implementation. We find a lack of simple digital and analog circuits that can be realized easily. We see that many of the work emphasized making the sensor node efficient in terms of power consumption, etc., rather than on the application of those nodes itself, which leads them to follow complex approaches for sensor node development, making development costly and time taking. Also, we find a minimal number of works utilizing the open-source tools and hardware for system development.

This paper presents the steps involved in developing a wireless sensor node using an open-source development board, Arduino Uno. We have designed units used for signal conditioning to make signals suitable for use as an input to the microcontroller. We have focused our approach to maximize the use of simple circuits like the Schmitt trigger in frequency measurements, RC filters, resistive voltage dividers,

and also readily available components like potential and current transformer and the open-source development board. With this approach, we reduce the cost and time of development.

The paper proceeds with a basic introduction to the wireless sensor node and the process of its development in the next section. The following section presents the results obtained on the implementation of the developed node. We have used this sensor node in a wireless sensor network prototype system developed to measure electrical parameters at a power distribution point. The last section concludes our work.

73.2 Development of Wireless Sensor Node

In this section, we present the steps followed in the development of the wireless sensor node. A wireless sensor node is a device that has sensors interfaced to a processing unit and a wireless transceiver connected to the processing unit for wireless communication, as shown in Fig. 73.1. This section starts with a discussion on selecting a microcontroller for the development of a sensor node. It then proceeds with the design and implementation of different sensor node units: the voltage measurement, current measurement, frequency measurement, and wireless communication units.

73.2.1 Microcontroller Selection

Going through the literature, we see that a wireless sensor node's general structure includes some sensors for measurement and a radio device for wireless communication and data transmission. This gives us an idea of the minimum requirements

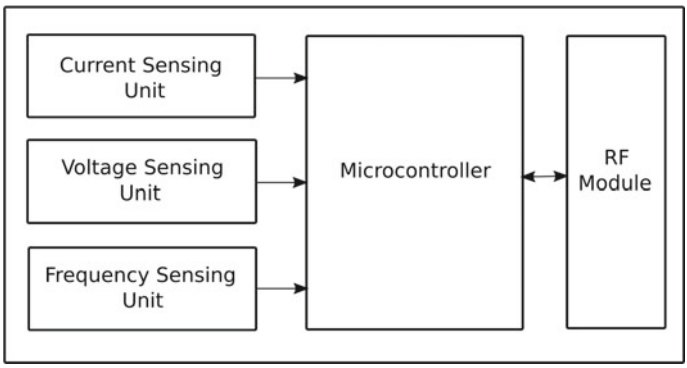


Fig. 73.1 Block diagram of proposed wireless sensor node

Table 73.1 Comparison of AT89c51 and ATmega328P microcontrollers

Features	AT89c51 (8051 Family)	ATmega328P (Arduino uno)
GPIO Pins	32	14 + 6 (Analog Input)
Built-In ADC	NA	6 channel ADC
Flash memory	4 kB	32 kB
On-Chip RAM	128 Bytes	2 kB
Timer/counters	2 16 bit	2 8-bit, 1 16-bit
Communication Ports	1 UART	1 UART 1 I2C 1 SPI
Clock speed	12 MHz	16 MHz
1 machine cycle	12 clock cycle	1 clock cycle

associated with the microcontroller like availability of a sufficient number of digital input/output pins, analog to digital converters, timers, and counters, communication ports, processing speed, etc. So, when we evaluate different microcontrollers and microprocessors available for embedded system development, we see that using a 16-bit or 32-bit processing unit in our application would be an overkill. Since we do not have to deal with large number's computations, there is no need for the high processing power of high bit depth microprocessors. This boils down our search to 8-bit processing units.

Based on the availability of development resources like development environments and hardware, development community base, and the microcontroller's learning curve, we select two 8 bit microcontrollers best suited for our purpose. A comparison of these two microcontrollers is made, as shown in Table 73.1, and ATmega328P is chosen for our work. The fact that Arduino Uno is an open-source development board based on our chosen microcontroller helps us to develop our sensor node fast utilizing the vast amount of technical documents already available. Also, it brings a universal appeal to our design as the work can be reproduced easily.

After selecting an 8-bit microcontroller for our node development, we proceed to develop other units of our sensor node.

73.2.2 Voltage Measurement Unit

To measure the voltage using a microcontroller, we need to condition the signal under observation, suitable for use as an input to the microcontroller [11–13]. The input range of the microcontroller we are using is 0–5 V DC, and the signal to be measured is around 230 V AC. So, we need to step down the voltage level using a step-down transformer to a safe working level. We use a 230 V/12-0-12 V potential transformer for this purpose. The reduced voltage obtained from the transformer is then rectified

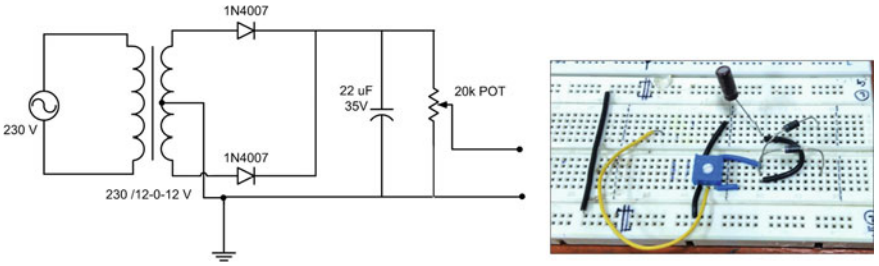


Fig. 73.2 Circuit diagram and breadboard implementation of voltage unit

using a full-wave rectifier. The rectified signal is then passed through an RC filter to get a smooth DC input signal, which we feed to the microcontroller's analog input, as shown in Fig. 73.2.

73.2.3 Current Measurement Unit

For the measurement of current, we will use a widely used Hall effect current sensor ACS712. It is available in three different ranges, and we select the lowest value sensor, i.e., 5 A range. With this sensor, the measurement of current up to 5 A can be easily done by using the current sensor directly. But that is not possible for high values of current. The current level is reduced using a current transformer similar to the approach adopted in the voltage measurement case to measure high current value. Here, we use a 30/5 A Tape Wound current transformer. The current obtained from the secondary is fed to the selected current sensor, and the sensor produces a voltage signal proportional to the applied current. This voltage signal is provided to the analog input of the microcontroller, as shown in Fig. 73.3.

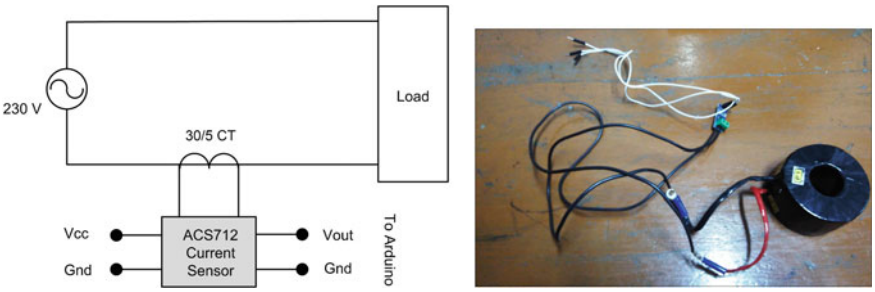


Fig. 73.3 Circuit diagram and hardware setup used in current unit

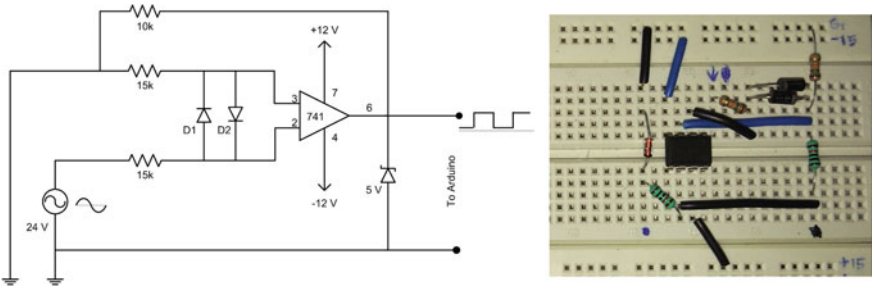


Fig. 73.4 Schmitt trigger circuit and its breadboard implementation

73.2.4 Frequency Measurement Unit

We have many methods available for the measurement of the frequency of alternating signals. The most commonly used techniques include using a zero-crossing detector circuit and a microcontroller that monitors and counts the events of zero-crossing in an alternating signal. The other method is converting the sinusoidal signal into square pulses and then measuring the pulse period using the microcontroller to determine the signal's frequency.

We have used the second method in our work, where we convert the sinusoid signal to a square wave using a Schmitt trigger, as shown in Fig. 73.4. This square wave is clipped using a Zener diode to discard the negative half of the square pulses, and the positive half is fed to the PWM pin of the microcontroller to measure the period.

73.2.5 Communication Unit

The wireless sensor node is called wireless because it is enabled to communicate using a wireless communication channel. For this purpose, we have selected the XBee radio module, which operates at a frequency of 2.4 GHz and works on the ZigBee standard, which is an open standard. These modules work in two different modes, AT mode and API mode, thus enabling us to have either a wireless serial link or a mesh network as per the need. We have used the AT mode of operation. The necessary configuration setting of the module is done as per the work [6], and then, the modules are connected to the serial port of the microcontroller as shown in Fig. 73.5.

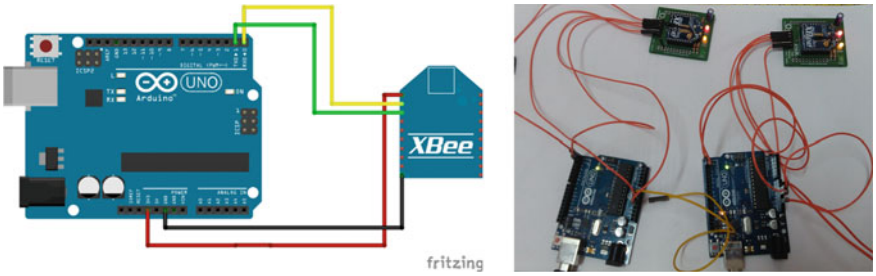


Fig. 73.5 Circuit diagram and implementation of Arduino Uno and XBee connection

73.3 Implementation of Proposed System

We have used the above developed wireless sensor node in our work, where we have implemented a power monitoring system using wireless sensor network. The power monitoring system is developed to monitor the basic electrical parameters at any distribution point in the distribution network within our institute campus. The system includes the above created wireless sensor node and the communication channel and data logger developed in our earlier work of WSN-based data logger [6].

The initial prototype is tested in the lab to validate the working of the complete system, as shown in Fig. 73.6. In this test, the sensor node takes electrical voltage measurements across the lamp load, the current drawn by the load, and the frequency of the applied voltage. On request from the central node, all these measurements are sent back to it using the ZigBee-based wireless serial link.

The existing monitoring system of our institute uses Elite 440 multi-line three phase meters made by Secure Meters Limited. These meters are good for operation in high tension feeder lines. But when it comes to their use in distribution lines, they are an overkill as we need to measure only few electrical parameters. Also, these

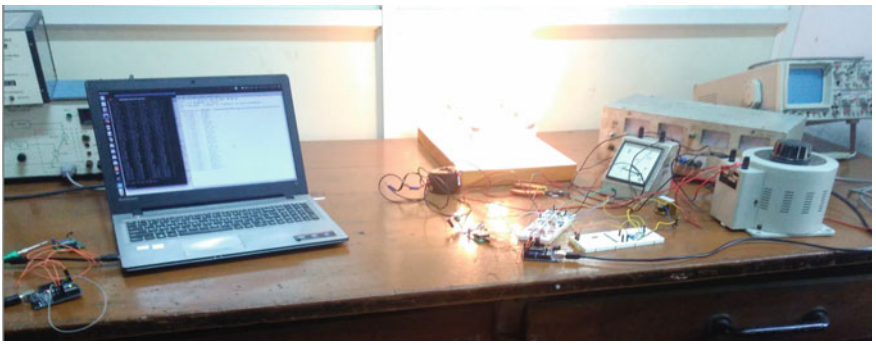


Fig. 73.6 Final setup of the power monitoring system prototype

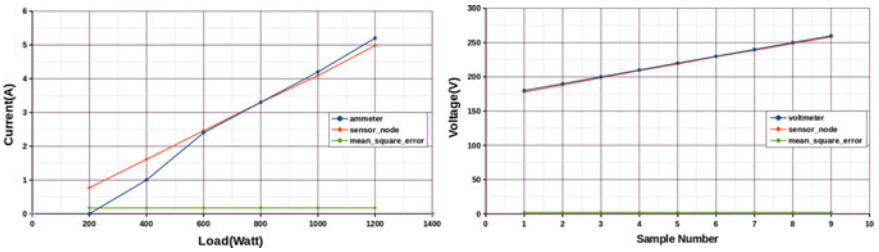


Fig. 73.7 Results of current measurement and voltage measurement

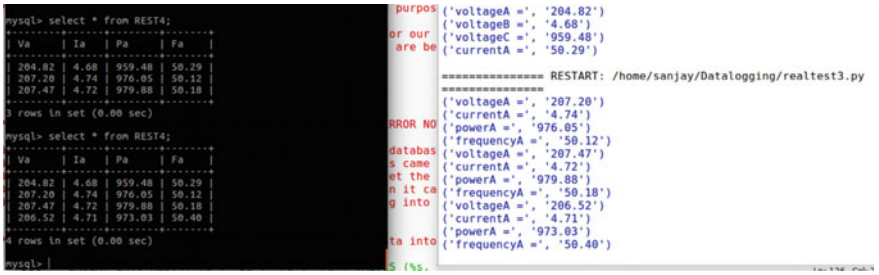


Fig. 73.8 Data loaded into the database and data received from wireless sensor node

meters are costly and are proprietary item leaving us with no scope of customization. Our proposed node was not only able to do accurate measurements with precision but was also very cheap compared to meters that are used. Also, the modular design of our wireless sensor node gives us the flexibility to customize them for specific application. The results obtained in our prototype testing were as expected, as shown in Figs. 73.7 and 73.8, and have validated the working of the wireless sensor node and the complete power monitoring system. The test outcome also validates our approach of using open-source hardware and software tools to develop a wireless sensor network. It also shows that the basic circuits being taught in theory can be realized in house using locally available components at low prices.

73.4 Conclusion and Future Scope

We made a wireless sensor node for our monitoring application. This sensor node measures three electrical parameters, voltage, current, and frequency of AC power. With this work, we have presented that a primary sensor node with good performance in terms of precision of measured values can be developed with limited resources and cost. The use of open-source resources like the Arduino Uno development board and basic circuits helps us focus more on the system's application than the system itself. In the future, we can optimize the work of microcontrollers to bring down the

unnecessary delays in processes running on the microcontroller. Also, we will work more on improving the signal conditioning units to minimize the noise levels in the signal. By doing this, we will have optimum performance from our sensor nodes.

References

1. Lee, D.S., Lee, Y.D., Chung, W.Y., Myllyla, R.: Vital sign monitoring system with life emergency event detection using wireless sensor network. In: SENSORS, 2006 IEEE, pp. 518–521. IEEE, New York (2006, October)
2. Yerra, R.V.P., Bharathi, A.K., Rajalakshmi, P., Desai, U.B.: WSN based power monitoring in smart grids. In: 2011 Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing, pp. 401–406. IEEE, New York (2011, December)
3. Cheong, P., Chang, K.F., Lai, Y.H., Ho, S.K., Sou, I.K., Tam, K.W.: A ZigBee-based wireless sensor network node for ultraviolet detection of flame. IEEE Trans. Ind. Electron. **58**(11), 5271–5277 (2011)
4. Han, D.M., Lim, J.H.: Smart home energy management system using IEEE 802.15. 4 and Zigbee. IEEE Trans. Consumer Electron. **56**(3), 1403–1410 (2010)
5. Korber, H.J., Wattar, H., Scholl, G.: Modular wireless real-time sensor/actuator network for factory automation applications. IEEE Trans. Ind. Inform. **3**(2), 111–119 (2007)
6. Chaturvedi, S., Parakh, A., Verma, H.K.: Development of a WSN based data logger for electrical power system. In: Gani, A., Das, P., Kharb, L., Chahal, D. (eds.) Information, Communication and Computing Technology. ICICCT 2019. Communications in Computer and Information Science, vol. 1025. Springer, Singapore (2019). https://doi.org/10.1007/978-981-15-1384-8_7
7. Yerra, R.V.P., Bharathi, A.K., Rajalakshmi, P., Desai, U.B.: WSN based power monitoring in smart grids. In: 2011 Seventh International Conference on Intelligent Sensors. Sensor Networks and Information Processing (ISSNIP), pp. 401–406. IEEE, New York (2011)
8. Suryadevara, N.K., Mukhopadhyay, S.C., Kelly, S.D.T., Gill, S.P.S.: WSN-based smart sensors and actuator for power management in intelligent buildings. IEEE/ASME Trans. Mechatron. **20**(2), 564–571 (2015)
9. Arora, J., Rawat, S.S.S., Srinivasan, K., Puri, V. et al.: Design and development of digital voltmeter using different techniques. In: 2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE), pp. 1–5. IEEE, New York (2014)
10. Ahamed, T., Sreedevi, A.: Design and development of pic microcontroller based 3 phase energy meter. Int. J. Innov. Res. Sci. Eng. Technol. **3** (2014)
11. John Errington's Experiments with an Arduino. http://www.skillbank.co.uk/arduino/measure.htm
12. Circuits4you. http://circuits4you.com/2016/05/13/arduino-ac-voltage/
13. All About Circuits. https://forum.allaboutcircuits.com/

Chapter 74
CommonKADS and Ontology Reasoner: Bulky-Baggage Case Study



Edwin Fabricio Lozada T., Wladimir L. Tenecota, Paul Santiago Pullas Tapia, and Livio Danilo Miniguano Miniguano

Abstract This paper presents the development of an agent based on an ontology reasoner based on the CommonKADS methodology. An ontology is built for the representation of the shoplifting trick "bulky-baggage" in the Protégé tool and in Java source code on the Jena library. And to build the agent, the Pellet reasoner is used both for Protégé and for its implementation in Java. As a result, the products that were stolen from the supermarket are taken by inference.

74.1 Introduction

Shoplifting or shoplifting refers to the theft of merchandise in stores during working hours. It is one of the most prevalent non-violent crimes in society, and the one that has increased the most in recent years. [1], this being a problem that costs customers and stores entrepreneurs billions of dollars per year since products tend to increase since entrepreneurs have to pay for lost merchandise, and their security and customers are obliged to purchase the products at higher costs. Shoplifters (shoplifters) use various tricks to steal merchandise from stores; one of them is the bulky-baggage. This trick

E. F. Lozada T. (✉)
Carrera de Software, Universidad Regional Autónoma de Los Andes, Ambato, Ecuador
e-mail: ua.edwinlozada@uniandes.edu.ec

W. L. Tenecota
Facultad de Ciencias Humanas de la Educación y Desarrollo Social, Universidad Tecnológica, Indoamérica, Ambato, Ecuador
e-mail: wladimirtenecota@uti.edu.ec

P. S. Pullas Tapia
Facultad de Ciencias Humanas y de la Educación, Universidad Técnica, Ambato, Ecuador
e-mail: paulspullas@uta.edu.ec

L. D. Miniguano Miniguano
Facultad de Ciencias Humanas de la Educación y Desarrollo Social, Instituto Superior Tecnológico Sucre, Quito Ecuado, Universidad Tecnológica, Indoamérica, Ambato, Ecuador
e-mail: dminiguano@tecnologicosucre.edu.ec

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_74

consists of the camouflage of the merchandise in larger packages. Packages that include gloves (gloves), newspapers (newspapers), packages (packages), and any item that can be carried in the hands and can camouflage the objects to be stolen. The packages are known as booster boxes and can even be gift-wrapped [2].

Video surveillance cameras provide a large amount of knowledge that can be used by knowledge engineering. The monitoring and detection of activities are very important in the security of a business. At present, it is possible for any user to connect surveillance cameras of their business and receive alarms on their mobile or email. In addition to giving them a greater sense of security, the automatic recognition tasks of human activities in video sequences are in high demand both by society and by industry and security agencies [3]. Therefore, to include ontologies in the procedures implemented by the agents, indirectly some reasoner is needed. Its main reason is that ontologies only explicitly describe a domain vocabulary, while reasoners obtain the implicit vocabulary hidden in them. [4, 5]. In this way, reasoners provide agents with a greater flow of information in order to improve the results of the objectives they pursue [6].

CommonKADS It is a methodology applied to the analysis and construction of knowledge-based systems (CBS). It consists of a set of models that allow various perspectives to be expressed from a given point that result in a set of activities for modeling. The models that comprise it are: Organization, Task, Agent, Knowledge, Communication, and Design. The CommonKADS methodology was developed and proposed by a group of researchers from various countries in the European community through a program called ESPRIT. And it is used in IC for the development of Systems Based on Knowledge (SBC) in a similar way to the methods used by Software Engineering. CommonKADS covers the complete software development cycle through a set of interrelated models that capture the main features of a system and its environment. The SBC development process is based on completing a set of templates of the models that are associated with states that represent the milestones in the development of each of them [7, 8]. The domain of this work is shoplifting under the buggy-baggage trick and aims to represent this knowledge in an ontology based on the CommonKADS methodology and to finally create an agent based on an ontology reasoner. The agent is built in Java with the Jena 2.6.4 libraries[1] with a Pellet reasoner 2.3.0[2]. This paper has been organized as follows in Sect. 74.2; the work methodology is presented. In Sect. 74.3, the experimentation is presented, and in Sect. 74.4 the conclusions, product of the experiment.

[1] https://jena.apache.org/

[2] https://mail.google.com/mail/u/0/?ui=2&view=bsp&ver=ohhl4rw8mbn4

74.2 Methodology

In accordance with the CommonKADS methodology for knowledge-based systems, the following models were developed.

74.2.1 Organization Model

See Tables 74.1 and 74.2.

74.3 Experimentation

To represent the knowledge of the buggy-baggage trick, an ontology called buggy-baggage was built in Protégé 3.4.5 for graphic display. For which the classes, properties and individuals of this profile are defined with their respective rules. In order to obtain the objects that were stolen, the following classes are defined: "Hidden_Objects" with the rule of obtaining all the products that disappeared in the second phase of a two-time period (on_Period_1, on_Period 2), "Shoplifter" which obtains the IDs of the customers who took one of the hidden objects and "Shoplifter-BuggyBaggage" which acquires the type of bulky package that the shoplifter used to hide the product.

Then Protégé's Pellet reasoner is used to see the inferences that occur in the buggy-baggage ontology, which allows to visualize the individuals that were inferred from the rules that were defined in the ontology.

As a result, we have in the class "Hidden_Object", "Shoplifter" and "Shoplifter-Buggy Baggage" the data of the objects that were hidden by one of the customers who entered the supermarket. Likewise, the Pellet reasoner is applied at the Java code level. As a result, the products that were stolen from the Supermarket by some of the customers are also presented. And what was the bulky bundle that the shoplifter used to hide the products. The following figure shows as results that customer_2 hid the products clothing_2 and shoes_2 in the bulky package buggy_baggage_2.

Table 74.1 Template OM-1: identification in the organization of problems and opportunities oriented to knowledge

Organization model	Problems and opportunities OM-1 worksheet	
Organizational context	Supermarkets are urban commercial establishments that sell consumer goods, including food, clothing, hygiene items, perfumery and cleaning They are stores that can be part of a chain, generally in the form of a franchise and can have more locations in the same city, state or country Their main characteristics are: They have their items on display so they are more prone to theft; they generally offer products at a low price. Also, their main objective is to generate profits so they try to counteract the low profit margin with a high sales volume	
SWOT matrix	*Weaknesses* The security techniques it has are not sufficient to counteract the high degree of theft they face. Clarke (2002) also tells us that people seem to have less inhibitions about shoplifting than they do about people. For they know that there is little probability that they will be discovered; And if they do, in most cases they can give some credible excuse like forgetting to pay *Threats* The articles and products they offer are on display so they are more prone to theft. Additionally, these sales methods are so com- mon today that they provide many opportunities for shoppers to manipulate merchandise and hide it in their clothing or in bags. (Clarke, 2002) *Strengths* They sell perishable and non-perishable consumer products so they are the favorites for customers *Opportunities* A high volume of sales generates greater profits	
Knowledge-intensive problems with the time factor as a determining factor in carrying out the process	One of the techniques that shoplifters (shoplifters) use to steal objects is bulky-baggage (bulky packages) that consists of camouflaging the articles or products in bulky packages So one of the main ones has to do with the fact of recording the specific knowledge of the tasks that shoplifters perform with the bulky-baggage trick. This will avoid the loss of large volumes of products due to theft and will allow the people who comment on it to be identified through the construction of an agent An effective response to the bulky-baggage trick of shoplifting means: fewer repeat offenders, decreased loss, increased sales, and increased profits	
List of problems with their associated priority	Taking into account the factor of time and economic resources as a determining factor, the following problems have been defined	
	Problems	*Priority*
	The constant theft of articles and products	High
	Theft of articles with the bulky-baggage trick	High
	Offer all your products and items on display	Low
	High economic investment for security personnel	High

Table 74.2 Worksheet CM-1: description of the transaction SEND ALERT

Agent: motion sensor	
Model agent	Agent leaf AM-1
Name	Motion sensor
	[Cognex Corporation (2011), mentions the PatMax technology that allows the location of objects to detect their movement]
Location in the organization	The motion sensor will be activated the moment a product begins to move, two sensors will be located in each supermarket stand
Type agent	It is a software agent that performs the process of detecting the movement of a product that will be being mobilized by a person
Restrictions temporary	The motion sensor agent should always be perceiving the movements of all supermarket products according to its stand
Communicates with	It will communicate with the task MPRO (movement of a product)
Knowledge	The sensing agent has knowledge of what can be recognized as a product and a person [López (2011) mentions some of the techniques used to carry out this activity, with the temporal derivative method]
Responsibilities	The motion sensing agent must detect the movements made by the product, which will be made by a person
Agent: ontology	
Model agent	Agent leaf AM-1
Name	Ontology
Location in the organization	The ontology has as entities the actions that a product can carry out, such as moving or hiding and by which person it will be carried out in order to infer if a product disappears in a scene and thereby determine who is the possible person who stole the product
Type agent	It is a software-type agent where all the entities involved in a product movement scene can be established
Restrictions temporary	The ontology must have the possible actions that a product can perform such as moving or hiding
Communicates with	The ontology agent will communicate with the task MPRO: movement of a product and LALR: issue alert
Knowledge	The ontology will have knowledge of the products that exist in the store, with the possible actions (move, hide) that it can take and the people who have the product in a certain period of time of the determined stand
Responsibilities	Track the movement of a product in a scene to be able to launch an alert when that product disappears from the scene
Communications model	Description of transaction worksheet CM-1
Transaction name/identifier	SEND ALERT
Information object	Broadcast an alert signal to the guards
Agents involved	Motion sensor + security personnel
Communication plan	See Fig. 6
Restrictions	In the event of an alert being generated, an alert signal is sent to the human agent
Information exchange specification	This operation is of the report type

```
Type: Hidden_Products
        Individual: clothing_2
        Individual: shoes_2

Type: Shoplifter
        Individual: customer_2

Type: ShoplifterBuggyBaggage
        Individual: buggy_baggage_2
BUILD SUCCESSFUL (total time: 6 seconds)
```

Code 1. The results of the hidden products of the "buggy-baggage" ontology inferred with the reasoning agent Pellet (Static Agent).

A dynamic agent is also built from the dynamic agent (Fig. 18) that allows us to know if a client is a shoplifter or not. Its operation is based on having the customer's data, the bulky package and the type of product that it takes as inputs; to output if this is a shoplifter; if the type of bulky package that I enter is a bulky-baggage for theft; and if the product is hidden or not.

74.4 Conclusions

CommomKADS is a comprehensive methodology, encompassing corporate knowledge management, engineering knowledge and analysis, the path to knowledge-intensive and application systems design. It offers theories, methods and scientific techniques to facilitate the representation of knowledge, we can model and understand the mental reasoning processes used by knowledge experts and reproduce them in a knowledge-based system, with a high focus at the engineering level. Protégé is a framework that has a set of tools for the construction of domain models and applications based on knowledge through ontologies, in a visual way, where the classes are organized in a hierarchy to represent the outstanding concepts of a domain, a set of spaces associated with the classes to describe their properties and relationships, and a set of instances of said individual classes of the concepts that contain specific values for their properties; You can also make inferences with reasoners that are installed as application plugins. It is Java-based, extensible, and provides a plug-and-play environment that makes it a flexible foundation for rapid prototyping and application development. Apache Jena is a Java framework to facilitate the creation of applications for the Semantic Web. It offers a collection of Java tools and libraries for development and linking with data from applications, tools and servers. It also includes an API for managing OWL ontologies and an inference rule based on a reasoning engine with RDF and OWL data sources. The library has Java classes to help developers write code to handle RDF, WL, and SPARQL with W3C recommendations. The inference engine is rule based to carry out reasoning based on OWL ontologies and storage strategies for saving RDF triples in memory or on disk.



Fig. 74.1 Ontology "buggy-baggage" in Protégé



Fig. 74.2 Ontology "buggy-baggage" Inferred with the reasoner Pellet in Protégé

References

1. Aloysius, J., Arora, A., Venkatesh, V.: Shoplifting in mobile checkout settings: cybercrime in retail stores. Inf. Technol. People (2019)
2. Elnahla, N.; From the Panopticon to Amazon go: an overview of surveillance in retailing. In: Proceedings of the Conference on Historical Analysis and Research in Marketing (2019)
3. L. Belova, «Experience of Artificial Intellience Implementation in Japan,» de *E3S Web of Conferences*, 2020.
4. Greco, L., Ritrovarto, P., Vento, M.: On the use of semantic technologies for video analysis. J. Ambient Intell. Human. Comput. (2020)
5. Gomez, H.F., Martínez-Tomás, R., Arias, S.A., Zamorano, M.R.: Using semantic technologies and the OSU ontology for modelling context and activities in multi-sensory surveillance systems. Int. J. Syst. Sci. **45**(4), 798–809 (2014)
6. Quan, Z., Haarslev, V.: A framework for parallelizing OWL classification in description logic reasoners (2019)
7. Zhigalov, K., Avetisyan, K., Markova, S.: Training video surveillance system for the purpose of object identification with the help of neural networks. In: IOP Conference Series: Materials Science and Engineering, vol. 537, no. 5 (2019)
8. Lulic, L., Stipanovic, B., Smoljic, H.: The role and representation of expert systems in making decisions on grating credit in banks in the republic of Croatia. In: Economic and Social Development: Book of Proceedings, pp. 289–303 (2019)

Chapter 75
Accelerated Sorting of Apples Based on Machine Learning



Kiron Deb, Shobhin Basu, Vamshi Krishna Palakurthy, and Nabarun Bhattacharyya

Abstract Accurate classification and grading are important for the quality assurance of apples since it is necessary to ensure that they retain their freshness and texture until they reach the end user following transportation and storage. We demonstrate that machine learning makes it feasible to separate high-quality apples from the poor-quality ones, which have signs of rotting, insect damage, and other forms of degradation. An alternative to the prevalent manual process of sorting and classification of apples is proposed. We focus on apples grown in orchards in Kashmir and USA. We analyze images taken following harvest using two machine learning models—an open source model as well as a well-known commercial tool. We also propose a hybrid approach for sorting perishable inventory that streamlines the process by using a high-throughput sorter based on image classification. We weed out the apples with surface blemishes in the sorting stage to identify potentially premium grade apples. The apples sorted by our approach are further screened by existing industry-standard procedures such as manual sorting or automated processes like acoustic analysis for classification into different grades. Using this two-step approach, we reduce the time for isolating the higher quality apples compared to a process that relies on a single-step sorting process only. This also achieves greater consistency and economies of scale.

75.1 Introduction

Automated classification and sorting of fruits and vegetables following their harvest have attracted a good amount of research interest in the recent past [1–3]. Most of the

K. Deb (✉)
Calcutta International School, Kolkata, India

S. Basu
Fremont High School, Sunnyvale, CA, USA

V. K. Palakurthy · N. Bhattacharyya
Centre for Development of Advanced Computing, Kolkata, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_75

sorting and grading, however, are still carried out manually. An efficient automated process would lead to greater consistency and economies of scale, and avoid problems with subjectivity. Food and groceries are low-margin, competitive businesses. Any improvement in determining fruit quality will reduce wastage during distribution and storage, which in turn will lead to higher profitability and customer satisfaction. For fruits like apples, which are considered a premium fruit for domestic consumption as well as export, a data-driven, consistent, and reliable sorting method is essential to gain a competitive advantage.

The criteria used to classify fruits vary from fruit to fruit. In the case of apples, both the appearance and the physical firmness are considered the two important criteria. Image analysis could serve as the first layer of quality assurance, which would discard apples that have poor appearance such as spots and blemishes which would turn off a prospective buyer. This is to be followed by a second round of inspection by manually feeling the apples. Alternatively, an automated process using acoustic sensors [4] could be used for this second stage. However, the initial screening should eliminate the need for second level of sorting for a large percentage of apples, if the goal is to select the premium grade of fruits.

For analyzing the apple images using machine learning, tools such as Amazon Sagemaker [5], Microsoft Azure [6] or Google Cloud AutoML Vision [7] could be used. However, for the purpose of this research, we wanted to share code using which other researchers can reproduce similar results with their own datasets. Hence, we decided to use ImageAI, an open source Python-based library for this purpose to train and validate our model. Google AutoML Vision is then used to validate our approach.

75.2 Description of Method

75.2.1 Training Data

We collected an image dataset consisting of 422 images of apples of different varieties, such as golden delicious, fuji, Washington red delicious, honeycrisp, as well as Kashmir variety. Each apple in this dataset has three images taken 120 degrees apart. This results in three images per apple. Many of these images were acquired using a cell phone camera with a resolution of 3024×2730. These images were annotated as “healthy” or “damaged.” As the attached image shows, the imperfections on the apple’s surface are easily visible to the naked eye.

75.2.2 Training Program

We adopted supervised learning as the process for producing a machine learning model. In the process of supervised learning, we feed a computer several data inputs along with the correct corresponding labels, so that it can infer patterns or correlations and generate a model that can predict the label of a previously unseen input with sufficiently high accuracy. In our case, the data inputs are the apple images, and the classification labels are "healthy" and "damaged."

We used ImageAI, a Python-based library that facilitates the training and implementation of computer vision models to help generate the model [8]. ImageAI uses as its backbone other libraries (dependencies): TensorFlow, OpenCV, Keras, and NumPy.

The underlying algorithm in the Python program for detecting the location of the apple in each image is YOLOv3 [9]. The YOLOv3 object detection algorithm is leveraged through transfer learning, wherein the knowledge gained in a pretrained YOLOv3 model is used as a starting point for training a new customized model in order to reduce training time and improve accuracy. By transfer learning through ImageAI, we produced a model that not only identified where each apple is in the image but also gave each detected apple a classification indicating its grade.

75.2.3 Testing Program

We feed the three images representing one apple into the image analysis software. The output is a message stating that the apple is damaged if any of the three images appears to have a spot or a blemish. In addition, the confidence score, which gives as a percentage the likelihood of the prediction being correct, is given as an output for each image. Representative outputs are given in Table 75.1. If the apple passes the image quality test, it proceeds for further testing to determine its firmness. Figure 75.2 shows the block diagram of the entire classification and sorting process.

75.3 Results and Analysis

We analyzed images of 32 apples from orchards of Kashmir obtained by Centre for Development of Advanced Computing (C-DAC), Kolkata using the above machine learning technique. Of these, we found 14 to be with some surface blemish, and the remaining 18 are eligible for further inspection, so that these can be certified as premium grade. Thus, 44% of apples could be eliminated right away by adding this image classification process. We expect this to lead to improved efficiency and sorting time.

Table 75.1 Apple images and corresponding model prediction

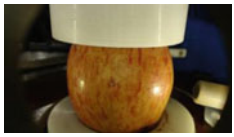
Image	Model prediction	Confidence score (%)
	healthy_apple	85.54
	healthy_apple	83.43
	healthy_apple	71.21
	damaged_apple	90.86
	damaged_apple	98.04



Fig. 75.1 Example image from our dataset consisting of 422 images

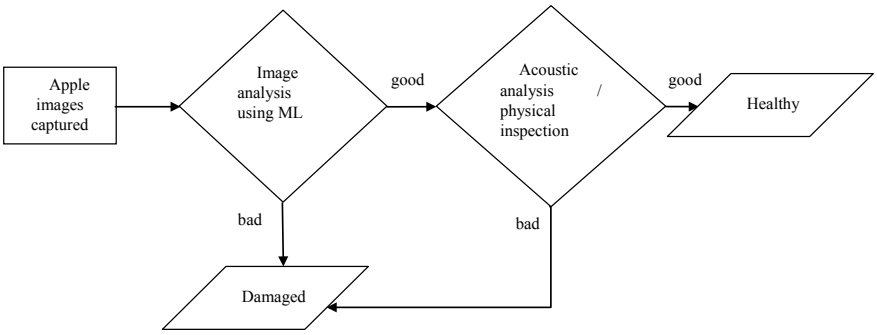


Fig. 75.2 Block diagram of the entire classification and sorting process

75.4 Validation Using Google Cloud AutoML Vision

As a validation of our approach, we used a commercially available machine learning product to predict apple quality. The product chosen for this purpose was Google Cloud AutoML Vision. There was no need to write code in this approach. By uploading the same dataset of apple images in the product's user interface, we were able to train two AutoML models. The first one is an object detection model that is capable of running in the cloud only. It was trained on a subset of 181 images where we drew the bounding box around the apples using the labeling tool in this product. The second one is a classification model that used the full dataset consisting of 422 images and is capable of running on edge devices. It can be exported from the product. Since YOLOv3 also has object detection as its model objective, a batch prediction request was sent for the test dataset of Kashmir apples to the AutoML object detection model hosted in Google Cloud. The predictions were similar to those shown in s with confidence scores typically over 95%, and lowest at 79%. Figure 75.3 shows the AutoML classification model that we trained on 422 images.

75.5 Further Work

Future enhancements could include developing a conveyor belt system to capture the apple images, and an automated analysis program using the machine learning algorithm. The conveyor belt system could be further integrated with a second level of automated system to test the firmness of the apples using non-destructive means such as acoustic sensing. Analysis of sound signals generated when the apple is physically tapped by a small hammer would identify any internal damage masked by a blemish-free surface appearance.

The apples that do not pass the image analysis test can be further screened for the degree of blemish, so that some of these apples can be sold, perhaps at a discount, if the magnitude of the blemish is reasonable. This will require further training of

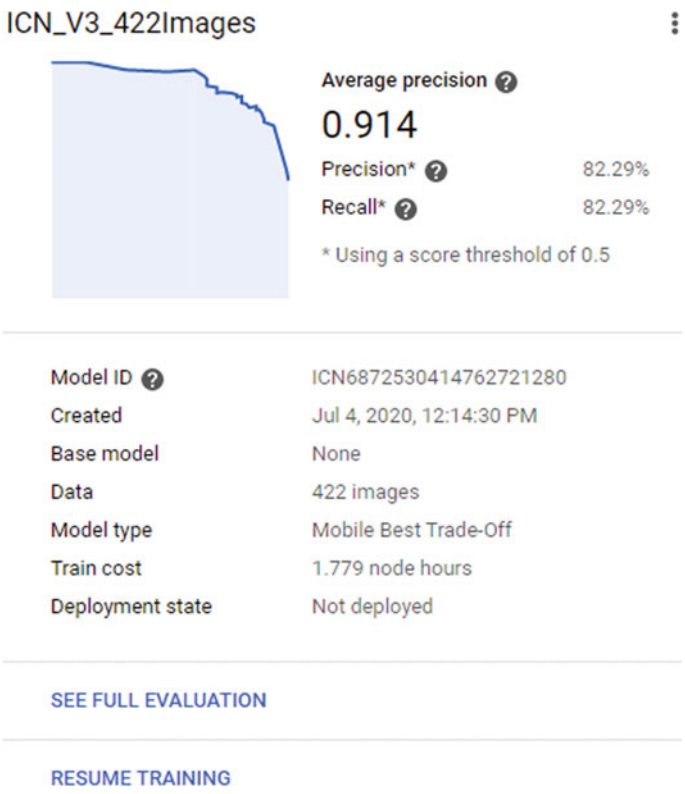


Fig. 75.3 AutoML Vision screenshot, adapted for style, shows the evaluation metrics from the classification model

the machine learning model since a high level of precision would be needed. For example, even though the third image given in Table 75.1 has a small spot, the model decided that the image indicated a healthy apple. However, the confidence score is relatively low at 71%, suggesting that though the apple crosses the threshold for healthiness, it is on the lower end and can hence be assigned a slightly lower grade. We intend to use the exported version of the AutoML classification edge model to build Web and mobile apps, so that interesting cases like these can be collected for training enhanced models.

75.6 Conclusion

We present a mechanism of non-destructive testing of freshly picked apples using machine learning. The proposed technique would serve as a first-level screen for separating the highest quality of apples based on their images alone. Initial tests

on a typical dataset of freshly picked apples indicate that up to 40–45% could be screened out if the goal is to identify superior apples free from any surface blemish. The tools used are open source software available at no cost. This is expected to lead to easier uptake at the beginning as no investment for software is required. Alternatively, commercial products like Google AutoML vision could be used as these maybe faster to implement and would require no coding effort.

Acknowledgements We have taken permission from competent authorities to use the images/data as given in the paper. In case of any dispute in the future, we shall be wholly responsible. Figure 3 is a section of the Google AutoML Vision UI. Google and the Google logo are registered trademarks of Google LLC, used with permission.

References

1. García-Ramos, F.J., Valero, C., Homer, I., Ortiz-Cañavate, J., Ruiz-Altisent, M.: Nondestructive fruit firmness sensors: a review. Span. J. Agric. Res. **3**(1), 61–73 (2005). https://doi.org/10.5424/sjar/2005031-125
2. Nourain, J.: Application of acoustic properties in non-destructive quality evaluation of agricultural products. Int. J. Eng. Technol. **2**(4) (2012). ISSN: 2049-3444
3. Abbott, J.A.: Quality measurement of fruits and vegetable, Elsevier. Postharvest Biol. Technol. **15**, 207–225 (1999). https://doi.org/10.1016/S0925-5214(98)00086-6
4. Abbott, J.A., Massie, D.R., Upchurch, B.L., Hruschka, W.R.: Nondestructive sonic firmness measurement of apples. Trans. Am. Soc. Agric. Eng. **38**(5), 1461–1466 (1995)
5. Amazon Sagemaker Image Classification. https://docs.aws.amazon.com/sagemaker/latest/dg/image-classification.html. Accessed 01 Sept 2020
6. Microsoft Azure Custom Vision. https://docs.microsoft.com/en-us/azure/cognitive-services/custom-vision-service/. Accessed 01 Sept 2020
7. AutoML Vision Beginner's Guide. https://cloud.google.com/vision/automl/docs/beginners-guide. Accessed 01 Sept 2020
8. Moses, O.J.: ImageAI, an open source python library built to empower developers to build applications and systems with self-contained. Comput. Vis. Capabilities (2018)
9. Redmon, J., Farhadi, A.: YOLOv3: an incremental improvement, 8 Apr 2018. arxiv.org/abs/1804.02767

Chapter 76
Execution of Smart Contacts Concurrently Using Fine-Grain Locking and Asynchronous Functions



Hemalatha Eedi and Pattan Asif Khan

Abstract Smart contracts are widely popularized to transfer an asset securely from one device to another. Smart contracts contain the blockchain's business logic that revolutionized the concept of self-enforcing contracts without the need for a middle man. Many use cases thus are being generated day by day. Smart contracts allow us to build decentralized applications through blockchains. The transactions in the smart contract are interrelated; they define the overall flow of the application. To maintain integrity between the transactions, the execution of the smart contracts is sequential and has no conflicts. These contracts' execution is slow and time-consuming. Thus, we cannot leverage the current systems' complete resource utilization. To overcome this, we propose a safer multi-threading model using fine-grain locking and thread executor service mechanism to concurrently execute the smart contract which also provides integrity among the transactions. Our results prove the faster and safer execution of smart contracts.

76.1 Introduction

Blockchain is essentially a decentralized distributed ledger, shared among a set of peers who communicate using a shared protocol. Each blockchain contains data stored in a format called "blocks," and each block contains a list of transactions, timestamps, and additional data required to execute the transactions. All the blocks are cryptographically linked to each other, i.e., each block is linked to the previous block using the previous block hash [1].

A smart contract is a code or a piece of a program, that can be used to transfer entities between two parties with strict monitoring, and these details cannot be manipulated by any means. Further, after the transfer, all the details are logged into the

H. Eedi · P. A. Khan (✉)
JNTUH College of Engineering Hyderabad, Hyderabad, Telangana, India

H. Eedi
e-mail: hemamorarjee@jntuh.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_76

blockchain to prevent any entity's illegal authorization. Two entities, "the miner" and "the validator," execute the smart contract. Miners are the nodes that own a block in the blockchain, and they execute all the transactions present in the block sequentially. Validators are the nodes that re-execute the transactions to verify the miner's identity, and on successful verification, they update their ledgers accordingly. Thus, all of the nodes in the chain maintain integrity in the ledger state. To ensure integrity among different nodes or to have the same state of the ledger in all the nodes, all the transactions are executed sequentially [2], resulting in the same state. This mechanism ensures the integrity and consistency of the state of the ledger across the system [3]. However, it fails to leverage the absolute power of multi-core machines today; also, the mechanism works slower to parallel mechanisms used by current database systems.

In this paper, we discuss mechanisms to ensure the state's integrity in all the systems, propose a concurrent system leveraging the power of multi-threading, improve the utilization of current computational systems, and achieve better results than the existing sequential mechanism using locking techniques [4]. Fine-grain locking, in this type of locking mechanism, rather than blocking the whole object and non-related processes by a thread [5] it concentrates on locking what is required, thus ensuring a higher level of concurrency, although with more number of locks, we have an over the head of maintenance, and it ensures security and a higher level of parallelism. Asynchronous functions refer to the routines, where there is no block scoping, i.e., one block of code does not depend on the other function or wait for other functions to execute the code [6]. All the tasks run parallelly in the background without interrupting the flow of each other. Fine-grain systems most implementations are based on the synchronous mechanism to ensure there is no race condition, but there is a possibility of relaxation for few functions to run asynchronously. We leverage this advantage and follow safe thread programming to achieve better results [7].

76.2 Related Work

In [8], the authors propose a method using lockless mechanisms and concurrent graphs; they use a software transactional memory (STM)-based methodology to ensure conflict serializability and integrity of all the transactions. STMs take care of most of the problems caused by concurrent access to data, crash failure, and a mechanism for recovery. In [9], the author proposed the idea of parallel blockchains, i.e., mining and validation on parallel blockchains. There can be discussed about updating, maintaining single history when operations performed on parallel blockchains; he discusses techniques such as transaction sharding, where transactions are distributed to pools of miners, and each pool can access a single shard, and a pool can only do execute transactions present in the shard. A reduced wait between block addition with a minimum number of conflicts is also added. In [10], in this paper, author proposed a locking-based mechanism, where he uses traditional locking methods to lock on

a resource, then execute the transaction, and record the locking procedure in the mining phase. Later in the validation phase, based on the record, a Happens Before graph is generated; all the transactions are executed through the graph and using a fork-join approach. The author proposes a parallel mechanism and raises issues about the data conflicts that could occur through the given mechanism. In [11], the author proposed the ACE (asynchronous and concurrent) execution mechanism, and he uses his infrastructure to have separate roles for users, miners, and service providers. Service providers are responsible for ordering and executing transactions. In [12], the author proposed a method where the state managements and smart contract executions are separated from the original blockchain insertions and updations; the author also mentions a separate pipeline model for verification and creation of blocks in parallel. In [13], the author proposed a transaction splitting algorithm for the execution of smart contracts. This helps in resolving the issues of synchronization while executing the smart contracts in parallel.

76.3 Proposed Methodology

The paper proposes the following methodology, to ensure concurrent execution of smart contract transactions. In the first step, a miner does parallel guessing of the nonce, using the brute force approach; nonce can be guessed parallel as the operation is self-isolated, and there is no dependency in guessing the nonce. At the end of this step, a miner would have successfully gained a block in the blockchain. After finding the nonce successfully, the miner who has won the block executes the smart contract; for example, we take a simple transfer and receive a smart contract. Each of the thread acquires a vertex lock on an entity present in the smart contract; here, an entity is referred to as the transaction's participants or an address present in the transaction.

During mining, all the instructions present in the node are executed atomically; atomic instructions are referred to the instructions that are executed fully, and only once, i.e., no two threads can acquire a lock simultaneously on the same instruction or resource. For each acquired lock on a transaction, we create a task and push it into the executor pool, and the threads acquire tasks from the executor pool through the blocking queue and execute the task. When all the transactions are executed, the program stops execution. Using the order of transaction execution, we generate a Happens Before Graph; a Happens Before graph is a directed acyclic graph (DAG) where nodes are the participating entities and edges consisting of relations between them. During validation, for each parent vertex in the Happens Before graph on which a lock is obtained. We try to acquire the child nodes or the entities to which the transaction must happen; we push it into a fork-join pool. All the statements are then executed by threads present in the pool.

The current paper works on permissionless blockchain like ethereum [14], where smart contracts are written in a language called solidity [15]. Solidity is a turing complete language and does not allow random execution of the transactions. Ethereum virtual machine (EVM) on which the code is executed runs on a single core

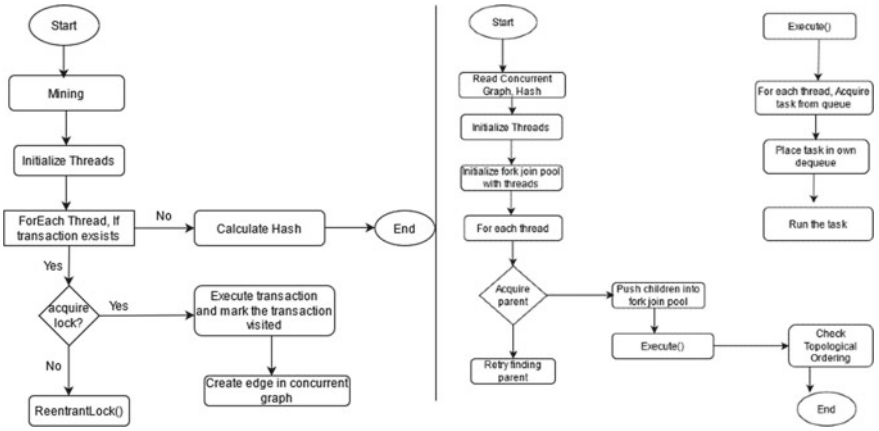


Fig. 76.1 Complete flow of mining and validation

system; this does not allow multi-threading or execution of transactions concurrently or in parallel, reducing the speed of executing the transactions.

To overcome this problem, the paper proposes to use Java, which is currently being used in permission blockchains, to write smart contracts and demonstrate the results using the same. Also, we do a comparative study based on the results obtained from both the systems. JVM has robust support for multi-threading and can handle threads well (Fig. 76.1).

In the mining process, a miner tries to guess the nonce to obtain a block in the chain to execute the smart contracts. Once a miner guesses the nonce correctly, he initializes a set of threads to execute the smart contract. Each thread in the transaction tries to acquire the lock on the objects present in a single transaction; if a lock is acquired, the transaction is executed successfully. Otherwise, a reentrant lock, i.e., the thread, tries to keep trying for a lock. As most transactions are distinct, the execution is done quickly. But in case there is a race condition in the reentrant lock, we can set a timeout for a thread after the thread retries exceed a certain threshold. Once the transaction is executed, the transaction is added as an edge into the concurrent graph. In the end, we generate the final hash of all executed transactions, which can be used for verification by the other peers present in the network.

In the validation process, the validators get a concurrent graph as input; the validator creates a set of threads; each thread finds a parent node and pushes all its children into a fork-join pool [16]. Each thread in the fork-join pool has its independent flow and acquires tasks from the queue; each thread pushes the tasks into its dequeue, using which all the tasks are executed. The threads which do not acquire tasks can barge in other tasks or acquire tasks from other threads. Thus, we must ensure no thread executes a transaction twice. Finally, after all the transactions are executed, a topological sort can be performed to verify whether all transactions are executed successfully, and the final hash generated matches the hash provided by the miner; if it does, block is accepted, else it is rejected.

```
Algorithm 1 MineParallel(T):

Input: Transactions present in the given smart contract

Output:- A Happens Before Graph(G)

       Final Hash of Transactions(H) and order of Execu-
tion

function MineParallel(T):

  forEach Transaction A in T:

    Obtain the lock on entities present in the Transac-
tion

      if(lock):

        Task.submit() //Submit Task into the thread pool

          ExecuteTransaction(A)

            if(vertex in G):

              addEdge(i,j)//CreateEdge in ConcurrentGraph

            Else:

              addVertex(A) //Createvertex if not found

              addEdge(i,j)

              hash()

              record()

      Else:

        ReentrantLock()

 Based on Locking log generate a Happens Before Graph(G)

 return (G,H)
```

```
Algorithm 2 ExecuteTransaction(A)

Input: Happens Before Graph

Output: Validation result

function ExecuteTransaction(A):

  For Transactions in Blocking queue:

    Obtain a transaction T from the queue

      If Transaction is Atomic:

        Execute The Transaction

          Update the state of atomic variables

            Else:

              Obtain lock on entities in the Transaction

                  if(lock):

                        Complete the task execution

                        Release the lock

                  else:  tryLock()
```

The concurrent miner uses the following technique to execute the transactions. For each transaction, say A in the transactions (T) list. A thread tries to obtain a lock on the entities present in the transaction. If it obtains a lock successfully on all the entities, it submits the transaction into the thread pool to execute the task. To make sure we do not execute a transaction multiple times, we can store each transaction data inside a ConcurrentHashMap() (H1), containing the list of already executed transactions. ConcurrentHashMap in Java is a thread-safe and also makes sure safe accessibility of data by the threads.

In the execution algorithm, whenever a task is submitted into the queue, the threads responsible for the execution of the task fetch a task. The blocking queue ensures that no two threads access the same resource. Hence, we do not have to worry about conflicts in such cases. For each task, we would need to acquire a lock. To know whether a particular entity or node is locked or it is not, we maintain a ConcurrentHashMap() which would contain the list of entities, and whether the entity is locked or not. After fetching the lock successfully, the task is executed by the thread. In case lock is not obtained, we do not do a thread.sleep() operation. If all threads are busy, then we might leave the core unused. To overcome that we use tryLock(), the concept of tryLock() in Java is that a thread tries continuously to obtain a lock on an object for a certain period and does not leave the core. Even after the mentioned time frame if a lock is not obtained, then the thread goes to sleep.

As in-memory operations are quicker and obtaining locks, checking for locks does not take significant time and is mostly O(1) operations. An equivalent of this, we could use a “mapping()” construct similar to a hash map for locking in smart contracts of

ethereum. The two operations, acquiring the locks and executing the transactions, can run in parallel, without conflicts, but, while we use locking techniques, the program flow restricts to either of the blocks, and the other threads do not perform any task, so while we submit tasks to the thread pool, they are all stored in a queue and wait for the execution.

```
Algorithm 3 Validator-Validator Function

Input - A Happens before Graph(G) and the result of
serial transactions

Output - The same Hash generated after the execution of
smart contract

function Validator(G):

      /*The graph execution starts from a parent nodes,
Parent node is a node with zero incoming edges, is non
dependant on any transactions*/

  For each parent node in G:

   if(getLock(Node parent):

      parent.[adjList].fork()

      parent..[adjList].join()

   Each Thread(Ti) does the following:

      Try to acquire a lock on a parent node,

      Push all the transactions list into Fork-Join pool

      Each thread in fork-join pool

          acquires a set of tasks Using Per-Queue threads

      Each Thread load tasks in local Queue and execute

      In case a thread does not acquire a task,
      It barges-in into local queues of other threads.

      After executing all the transactions call join()

      Return result
```

The advantage of the fork-join procedure in this process is that fork-join pools mostly will not allow tasks that cause conflicts to execute concurrently, and fork-join pools work well in the case of smart contracts because all the statements are mostly dependent and are short, thus providing us with a safer state at the end of the all join operations. In the end, we can compare both states of validator and miner if both reach the same state. Thus, both of them are having a good history, and we can update the history locally.

76.3.1 Implementation in Java

We use the concurrent package present in Java to generate the required number of threads using the FixedThreadPool class to avoid concurrency-based issues. The number of threads running will always be equal to the number of cores present in the system since Java implements OS-level threads that take a core of the system to run the thread. All the threads that do not acquire a task will be waiting until they are assigned with the task. All the threads which have acquired a task work along with the CompletableFuture class to execute the tasks asynchronously, i.e., task execution does not restrict the flow in a single block and gives us the facility to keep the main thread running. CompletableFuture implements a call-back-based mechanism to execute the task and can run in the background without disturbing the overall flow. We use AtomicInteger class to store data as it restricts access to the same variable by threads simultaneously and provides safe access to variables.

In the validation process, we create a fork-join pool, each thread in the pool has its dequeue where it stores the tasks, and to acquire a task, fork() is called. The threads then execute the tasks which are acquired via fork and are present in their dequeue. Here, some of the threads might get fewer tasks and can complete their execution fastly. In this case, the barge-in operation occurs; each thread tries to acquire tasks from the other queue and execute the tasks. This operation can happen because all of them use a similar shared memory. The barge-in operation improves the speed significantly. But, it comes with the overhead to check whether a task is previously executed, this problem of lookup can be solved using the concurrent hash map where we can keep a record of transactions which are already executed. At the end of this phase, we can perform a topological ordering of the graph to check whether all the transactions are executed or not.

76.3.2 Smart Contracts Used for Results

We use the Ballot.sol, SimpleAuction.sol, and TransferContract.sol smart contracts to test our methodology. A ballot is a smart contract that represents an electronic voting system. It consists of functions as a vote() and WinningProposal() functions similar to that of regular voting systems. It also has added functions such as delegate() and RighttoVote(), which can be used to transfer one's vote and to check whether a voter is a valid voter or not.

It is a smart contract where multiple bidders try to bid for an item, and there is a single owner of the contract. The owner initiates the bidding process. Each user who bids has his specific address given along with the bid amount; the bid takes place using the bid() function of the contract; the bidding amount and the bidder's address are stored in a "mapping." Finally, except for the winner, all the amounts of other bidders are reimbursed using the withdraw() function.

Table 76.1 Average speedup times of the contracts compared to as of serial execution

Operation	Ballot	Simple auction	Transfer
Mining	1.71x	1.32x	1.59x
Validation	1.83x	1.6x	1.9x

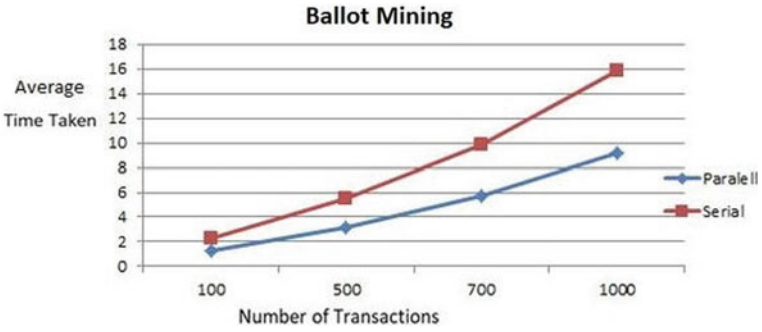


Fig. 76.2 Ballot mining

It is a smart contract used to transfer an amount or an entity from one address to another. It consists of method transfer(), using which we can transfer an asset from one address to another receive (); this method's functionality is to acquire an asset. Check() function to check whether an asset is received or not. A "mapping" consists of details of users between whom the transfer is going to happen. After the execution of this smart contract, the successful transfer of assets is completed.

76.4 Experimental Results

We run our simulations on an Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz with octave cores and 16 GB RAM. The codes run on the windows 64-bit operating system using Java libraries. We carried out executions using a randomly generated graph with varied nodes. Each smart contract consists of a minimum of 700 transactions. The transactions go up to 2000 per contract. The input taken is in the format that suits the need of the contract. The average speed up the time of the contract execution is given in Table 76.1. When compare to serial execution, our method improves the speed as shown in Figs. 76.2, 76.3, 76.4, 76.5, 76.6, and 76.7.

76.5 Conclusion and Future Scope

Finally, we can conclude that the methods proposed above bring a faster execution of the smart contracts than the serial execution; we can select a set of functions

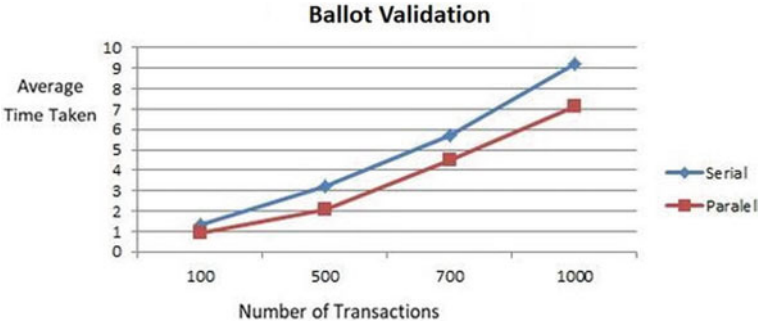


Fig. 76.3 Ballot validation



Fig. 76.4 Transfer mining

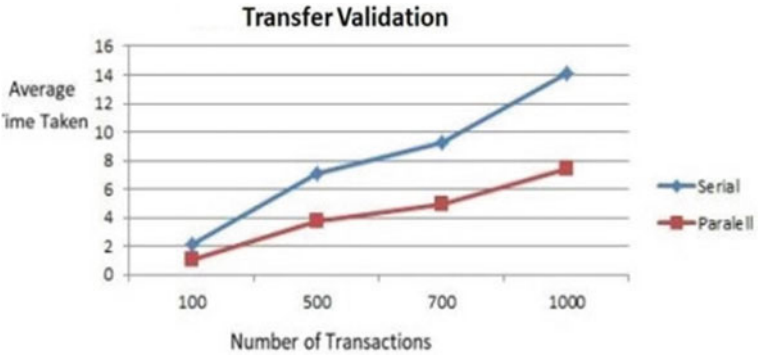


Fig. 76.5 Transfer validation

that can run asynchronously or tasks that can be allowed to run asynchronously and improve the performance significantly. Although the method comes up with the advantage of faster execution, maintaining the locks is an overhead and can increase the program's gas cost. There is also a problem that can optimize the execution further

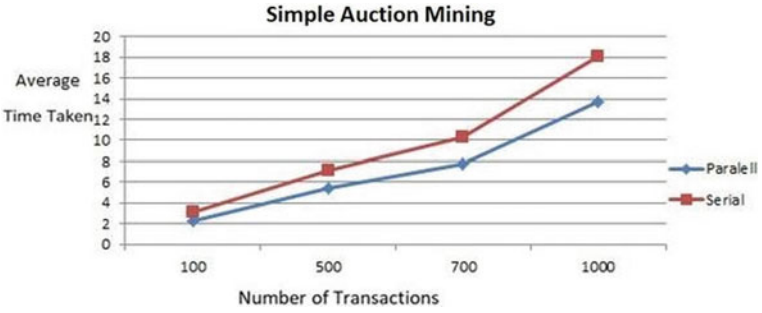


Fig. 76.6 Auction mining

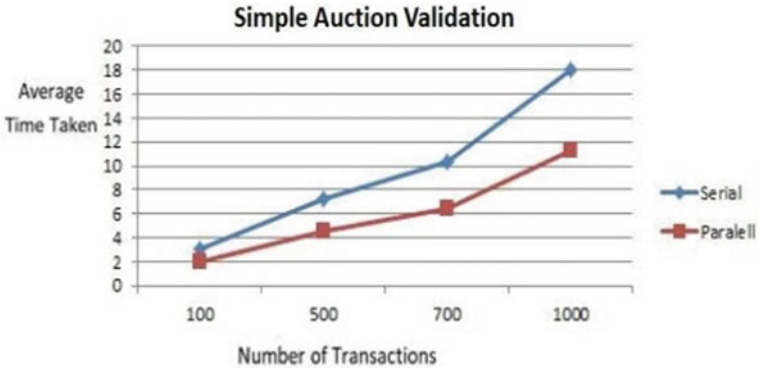


Fig. 76.7 Auction validation

which is to select which function to run asynchronously, and selecting these functions is sometimes hard and can require a complete understanding of the smart contract code. In the future, advanced optimization techniques can improve task assignments to the threads in the fork-join pool.

References

1. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system May 2009
2. Herlihy, M.P., Wing, J.M.: Linearizability: a correctness condition for concurrent objects. ACM Trans. Program. Lang. Syst. (TOPLAS) **12**(3), 463–492 (1990)
3. Flanagan, C., Abadi, M.: Types for safe locking. In: European Symposium on Programming, pp. 91–108. Springer, Berlin, Heidelberg (1999)
4. Szabo, N.: Formalizing and securing relationships on public networks. First Monday (1997)
5. Dice, D., Shalev, O., Shavit, N.: Transactional locking II. In: International Symposium on Distributed Computing, pp. 194–208. Springer, Berlin, Heidelberg (2006)
6. Kleiman, S., Shah, D., Smaalders, B.: Programming with threads. Sun Soft Press, Mountain View (1996)

7. Oaks, S., Wong, H.: Java Threads: Understanding and Mastering Concurrent Programming. O'Reilly Media, Inc. (2004)
8. Anjana, P.S., Kumari, S., Peri, S., Rathor, S., Somani, A.: An efficient framework for optimistic concurrent execution of smart contracts. In: 2019 27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), pp. 83–92. IEEE (2019)
9. Fitzi, M., Gazi, P., Kiayias, A., Russell, A.: Parallel chains: improving throughput and latency of blockchain protocols via parallel composition. IACR Cryptol. ePrint Arch, 1119. (2018). https://eprint.iacr.org/2018/1119
10. Dickerson, T., Gazzillo, P., Herlihy, M., Koskinen, E.: Adding concurrency to smart contracts. Distrib. Comput. 1–17 (2019). https://doi.org/10.1145/3087801.3087835
11. Wüst, K., Matetic, S., Egli, S., Kostiainen, K., Capkun, S.: ACE: asynchronous and concurrent execution of complex smart contracts. IACR Cryptol. EPrint Arch. **2019**, 835 (2019)
12. Yu, L., Tsai, W.T., Li, G., Yao, Y., Hu, C., Deng, E.: Smart-contract execution with concurrent block building. In: 2017 IEEE Symposium on Service-Oriented System Engineering (SOSE), pp. 160–167. IEEE (2017)
13. Yu, W., Luo, K., Ding, Y., You, G., Hu, K.: A parallel smart contract model. In: Proceedings of the 2018 International Conference on Machine Learning and Machine Intelligence, pp. 72–77 (2018). https://doi.org/10.1145/3278312.3278321
14. Ethereum. https://github.com/ethereum/
15. Solidity https://solidity.readthedocs.io/
16. Naftalin, M.: Mastering Lambdas: Java Programming in a Multicore World. McGraw-Hill Education Group (2014)

Chapter 77
Artificial Intelligence-Based Automatic Evaluation Engine



Sai Gayatri Vadali and Gali Suresh Reddy

Abstract Automatic evaluation of transcripts deals with comparing texts and evaluating them using the best available methods in NLP. The core part of this project is to enable the machine to evaluate descriptive transcripts all by itself. This reduces the human effort correcting huge answers in educational institutions and crowd examination platforms. Finding syntax and semantics out of text has been of great effort since years. As a part of this research, many NLP models like entity recognition, POS tagging, and topic modeling like LSA and LDA have come into existence. We are using a combination of these to get the semantics out of a transcript to eventually evaluate it.

77.1 Introduction

Natural language processing in answer evaluation has been a great research domain since years. Various NLP methods have been proposed in the literature to reduce blurring effects. Many of these methods are based on preprocessing and post-processing data. These techniques help us build machines which can understand natural language. In other words, natural language processing enhances the analysis on text data by removing the noise and the inconsistencies of the text. Natural language processing consists of different techniques, i.e., tokenizing [1], POS tagging, parsing, semantic similarity [2], and language model. NLTK [3] provides tool for above method. Its methods can be broadly classified into two that is preprocessing and modeling stage. POS tagging, parsing, topic modeling [4], WordNet-based similarity [5, 6], and Word2vec [7] are special domain methods.

S. G. Vadali · G. Suresh Reddy (✉)
VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, Telangana, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_77

77.1.1 Motivation

Auto-evaluation engine helps in evaluating the descriptive transcripts with high speed, accuracy, and transparency. This kind of a tool can help evaluators, teachers, instructors, and interviewers by reducing the time needed for evaluation and by giving them the basic understanding of student's capabilities even before they start evaluating the transcripts.

There has been a lot of research in this area since the beginning of natural language processing and text processing. Using the conventional technologies together with recent advancements in deep learning, a novel approach has been developed to build the engine.

77.1.2 Relevant Work

Nevertheless, automating descriptive answers can be particularly difficult because they cannot be restricted to certain rules. Also, getting semantics out of text can be very difficult. Although there are parsers like that of by Stanford dependency parsing [8] for grammar checking, certain problems like dealing complex sentences were not solved. Getting semantics using Word2vec [9] also did not solve the problem of correcting answers accurately.

As our approach is one such tries evaluate open-ended answers for generic essays, we have come up with a rule-based parser and WordNet-based and topic modeling-based semantic extraction methods.

We use a novel approach of making syntactic check of answer followed by semantic check. These two together give an overall analysis of understanding and expressive ability of the candidate in writing essays.

Following research works have already been made in NLP relevant to grammar check and semantic extraction which have been used in our present work. Ng et al., describe latent Dirichlet allocation (LDA) [10], a generative probabilistic model for collections of discrete data such as text corpora. Wenli [11], the basic principle of classic traditional information retrieval model, is the machine matching of the key word, namely retrieval based on keywords.

Also, VADER sentiment analysis [12] has been used for sentiment analysis at the end of semantic extraction. The idea of text summarization has also been included when converting bigger paragraphs of suggested answers to smaller answers. Character-level text classification has been studied for punctuation analysis. LSTM based question and answering model has been analyzed [13]. Grammar checking tool is based on deterministic grammar generation.

77.2 Background Work

77.2.1 Spell Check

A passage full of spelling mistakes may become very difficult to read and understand. A single letter change in the word may lead to change in the meaning of whole of the sentence, for example, consider the words bear and bare. With a small change in placement of letters, the meaning has changed totally.

77.2.2 Grammar Check

In natural languages like English, just like programming languages, following grammar rules are very essential to build understandable sentences. These rules form Syntax of the language. Syntax can be built on sentence structure like noun, verb phrases (rule-based parsers) or with the dependency of the rest of a sentence on a word (Stanford Parser). Techniques like CFG have been used for this purpose. Context-free grammar known as CFG helps in building a rule-based grammar parser. Context-free grammar-based tree can be parsed either top-down or bottom-up approaches. CFG is based on building tree with non-terminals and terminals. Non-terminals in natural language constitute NP, VP, and N which are finite. But terminals which are nouns like Joy and verbs like come are infinite. This poses challenge to rule-based parsers.

77.2.3 Semantic Extraction

Semantic extraction helps in extracting the meaning and understanding from a given paragraph. Techniques like topic modeling can help in extracting the semantics.

77.2.3.1 Topic Modeling

Topic modeling is the methodology in which topic of the paragraph is extracted. It extracts not only a single topic but then topics which constitute important part of the passage. Topic modeling has become famous by the techniques like LDA, LSI. LDA is based on the most usage of the word, and LSI is based on semantic indexing.

77.2.3.2 Sentiment Analysis

Sentiment analysis can help us understand the opinion of a person with respect to a topic. This helps us understand the personality and emotional ability of the person. Particularly there can be questions like

Do you think growth in investment banking—a sign of growth of nation?

Questions like this give a clear understanding of person's opinions, weightage to his/her points, analytical, and persuasive capabilities, and sentiment analysis helps us just do that.

77.3 Problems Posed by Existing Systems

77.3.1 No Generic Grammar Checker

English is a diversified language with many different grammar rules. The syntactic engines present right now cannot handle all the different types of sentences effectively. For each sentence, a separate grammar must be defined.

77.3.2 Extracting Context or Semantics of the Paragraph

The existing methods cannot correctly find the context of the sentences. Semantics has mostly been restricted to getting meaning from dictionary. The present method is based on vector embeddings and hence not totally reliable. Synonyms cannot be handled properly.

- I went fishing for some sea bass.
- The bass line of the song is too weak.
- Here 'bass' has different meanings in different sentences.

77.3.3 Dealing with Variety of Sentences

Complex sentences have different sentences in same sentence. Present parsers cannot identify the work put in making good complex sentences. Handling complex sentences itself is a difficult thing.

77.3.4 Humor and Sentiment Analysis

Humor and sentiment analysis can be difficult for a machine to understand. Often, understanding semantics is a very difficult task. A fascinating series of experiments sheds light on how humor affects the brain's functioning. When horizontal moving lines are presented to the left visual field (processed by the right hemisphere), and vertical moving lines are presented to the right visual field (processed by the left hemisphere) subjects report seeing either vertical stripes or horizontal stripes, but only very rarely both (a "crosshatch" of intersecting lines). This is called "binocular rivalry"—meaning only one image makes it to our awareness at a time, and the other is not consciously noticed. This is missing with the machines. Hence, humor becomes difficult.

77.3.5 Time and Space Complexity

The present system has difficulties handling time and space complexities. When the sentences are large that is, if the sentence is complex or have many numbers of words, the parser generally takes lot of time to generate parse tree. If that is the case, the system may go into halt state. If it is recursive descent parser, it goes into infinite recursion.

77.4 Proposed Work

The answer written by the candidate undergoes syntax check, semantic check followed by sentiment analysis. Here is a case study of it that occurs.

Let us try to see how the engine works for answer written to the given question "Write about Corruption." The answer written by the candidate in the text box undergoes following steps.

77.4.1 Preprocessing of Text

Text processing occurs at two times—before syntax check and before semantic check. Before syntax check, the passage is divided into sentences. and POS tagging is applied. Before semantic check, techniques like removal of stop words and punctuation are applied.

- Parts of speech (POS) tagging helps us to classify words of sentence into nouns, pronouns, and such English grammar parts of speech. We used Penn Treebank

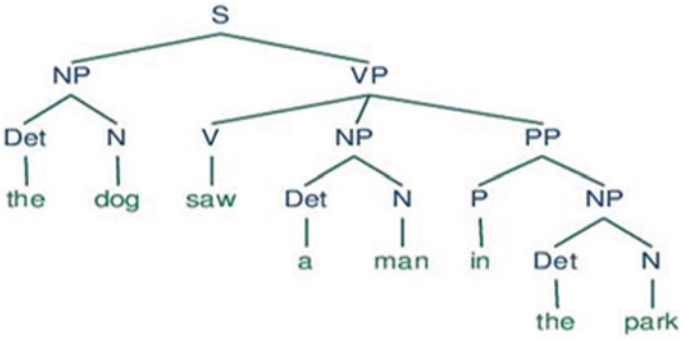


Fig. 77.1 Parse tree generated syntax rules

POS tag set for tagging. It has 36 tags in total. It includes different kinds of nouns, verbs, prepositions, punctuation, adverbs, etc.

- Chunking is the process of joining the words into phrases like noun and verb phrases. Chunking helps to join words and parse them through the rules. It is one of the crucial steps in text preprocessing.

77.4.2 Syntax Analysis

Syntax analysis is a process where the preprocessed sentence is parsed through a parse tree. The parse tree is generated based on syntax rules that have been formulated earlier. If a sentence is completely parsed through the syntax tree, we can say that the sentence is grammatically correct. If not, it has some grammar issues. Following is one example of the same (Fig. 77.1).

77.4.3 Semantic Analysis

Topic modeling is a process of extracting important topics from text. We find the topic from the question as well as from the answer written by the candidate. If they match, then the answer is with in context of the question. LSI and LDA are the topic modeling techniques used in our model.

77.4.4 Sentiment Analysis

The next step is to apply sentiment analysis to find out the sentiment of the text. This acts as an additional feature to find out the characteristics of the text.

77.4.5 Final Score

Final score is obtained by taking scores from all the above steps. For the answer in "corruption," following score is obtained. The scores obtained for the answer are as follows.

Syntactic score—75
Semantic score—41
Sentiment—0.89

We observe the scores are very satisfactory.

77.5 Experimentation and Results

The parser has been used to parse different types of answers written by the candidate as well as questions that can be asked.

77.5.1 Different Types of Answers that Can Be Written by Candidate

Following are the various ways in which a person can write the answer:

1. Grammatically correct answer with good semantic match.
2. Answer written by the candidate is gibberish and irrelevant to question.
3. Grammatically correct sentence but no semantic match.

One example is illustrated below with pictures. Question based on corruption was asked, and suggested answer was given. Then, the answer written by a good candidate is given. Refer Fig. 77.2 for answer by a good candidate and Fig. 77.3 for suggested answer.

A good syntactic and semantic score of 100 and 65 have been given, respectively, by the machine for above answer.

Corruption is misuse of authority to earn monetary gains in wrong means. Corruption is highly prevalent in many developed and developing nations. It can be controlled only with right governance and attitude of people. Politicians, government officials should not yield to corrupted sources of income. Huge penalties should be imposed on people who resort to unethical corrupted practices.

Fig. 77.2 Answer by a good candidate

As we all know that corruption is very bad thing. It inhibits the individual growth as well as society and country growth and development. It is social evil which is playing humans body and mind socially, economically and intellectually. It is continuously making its roots so deeply because of the increasing human greediness towards money, power and position. Corruption is the misuse of authority, public position, natural or public resources, power, etc. by someone to gain his/her personal gratifications. According to the sources, it has been identified that India ranks three in the highly corrupted countries. Corruption is highly spread in the field of civil service, politics, business and other illegal fields. India is a famous country for its democracy but it is corruption which disturbs its democratic system. Politicians are highly responsible for all type of corruption in the country. We chose our leaders by having lots of expectations to them to lead our country in the right direction. In the starting they make us lots of promises however, just after the voting they forget all that and involve in corruption. We are sure that our India would be corruption free a day when our political leaders would be free of greediness and use their power, money, status and position in right direction to lead the country, not their own luxury and personal wishes.

We should select very honest and trustworthy leaders to lead our India just like our earlier Indian leaders such as Lal Bahadur Shastri, Sardar Vallabh Bhai Patel, etc. Only such political leaders can reduce and finally end the corruption from India. Youths of the country should also need to be aware of all the reasons of corruption and get together to solve it in group. Increasing level of the corruption needs to take some heavy steps to get control over

Fig. 77.3 Suggested answer

77.6 Conclusions and Future Work

Writing has been the source of our knowledge as well as a powerful way of expressing our ideas in a way which stays forever. Also, a beautifully written essay can depict the personality as well as creative power of the individual. Auto-evaluation engine with these features of grammar checker, semantic analysis, and sentiment extraction can be a very good tool to assess as well as preserve the written wealth of the people. The present tool can be extended into an advanced version by using deep learning and generative language modeling for semantic analysis. Probabilistic methods like PCFG and generative methods like language modeling can help in revolutionizing syntactic checking of the text. Generative methods enhance semantic check to a great extent.

References

1. Vijayarani, S.: Adv. Comput. Intell. Int. J. (ACII) **3**(1) (2016)
2. Meng, L., Huang, R., Gu, J.: A review of semantic similarity measures in wordnet. Int. J. Hybrid Inf. Technol. **6**(1) (2013)
3. Manning, C.D., Surdeanu, M., Bauer, J.: The stanford core NLP natural language processing toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55–60. Baltimore, Maryland USA, June 23–24 (2014)
4. Jelodhar, H., Wagi, Y., Yuan, C.: Latent Dirichlet allocation (lda) and topic modelling, applications, survey. Nanjing University of Science and Technology (2018)
5. Pedersen, T.: WordNet: similarity—measuring the relatedness of concepts. In: Proceedings of the 19th national conference on Artificial intelligence, pp. 1024–1025 (2004)
6. Miller, G.A.: WordNet: a lexical database for English. Commun. ACM **38**(11) (1995)
7. Zhang, Y., Ma, L.: UsingWord2Vec to process big text data. In: Big Data (Big Data), 2015 IEEE International Conference on 29 Oct–1 Nov (2015)
8. Chen, D.: Christopher D Manning. Stanford University, A fast and accurate Dependency Parser using neural networks (2014)
9. Víta, V.K.: Word2vecbased system for recognizing partial textual entailment. In: Computer Science and Information Systems (FedCSIS), 2016 Federated Conference on 14 Sept (2016)
10. Blei, D.M., Andrew, Jordan, M.I.: Latent Dirichlet allocation. University of California, Berkeley. Berkeley, C A 94720. J. Mach. Learn. Res. **3**, 993–1022 (2003)
11. Deshmukh, A., Hegde, G.: A literature survey on latent semantic indexing. Int. J. Eng. Invent. **1**(4), 01–05 (2012). ISSN: 2278-461
12. Hutto, C.J., Gilbert, E.: VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM (2014)
13. Wang, D., Nyberg, E.: A long short-term memory model for answer sentence selection in question answering. In: 7th International Joint Conference on Natural Language Processing, Association for Computational Linguistics, pp. 707–712, July (1995)

Chapter 78
SmartRPL: Secure RPL Routing Protocol for Internet of Things Using Smart Contracts



P. Subhash, P. Navya Sree, and M. Pratyusha

Abstract Numerous nodes are connected to Internet of Things (IoT), nodes need to communicate and synchronize with each other, and these nodes are resource constrained with limited processing power, memory, energy with unstable links, and low bandwidth. There is a standard routing protocol for IoT-LLN (low-power and Lossy networks) known as RPL (routing protocol for low-power and Lossy networks); this protocol is used for communication between low power devices and exchange data. For secure communication among IoT nodes, the routing protocols require central authority which results in various security attacks and can severely degrades the network performance. In the proposed system, the blockchain-based framework with smart contracts is used to offer secure route establishment between IoT nodes in the network and protects the control packets information from being modified by an external attacker. Thus, it can prevent the formation of variety of attacks and to ensure secure route formation between nodes in the network.

78.1 Introduction

Routing protocol defines how nodes communicate and with one another and distribute data which allows them to select the route between the nodes [1]. Routing in IoT nodes is performed using RPL which is a proactive routing protocol for LLN such as IoT; these devices in the network are arranged in the form of directed acyclic graphs (DAGs) which has multiple destination-oriented directed acyclic graphs (DODAGs); each DODAG is in the tree-based structure which is created by the root node; these nodes are arranged such that no loops are formed in the network. RPL protocol is effective and secure routing protocol. These messages are carried via ICMPV6 messages and use the DODAG information object (DIO), DODAG information solicitation (DIS), and destination advertisement object (DAO) control messages used in RPL protocol [2]. In addition to this, offering security during the path establishment

P. Subhash (✉) · P. Navya Sree · M. Pratyusha
VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_78

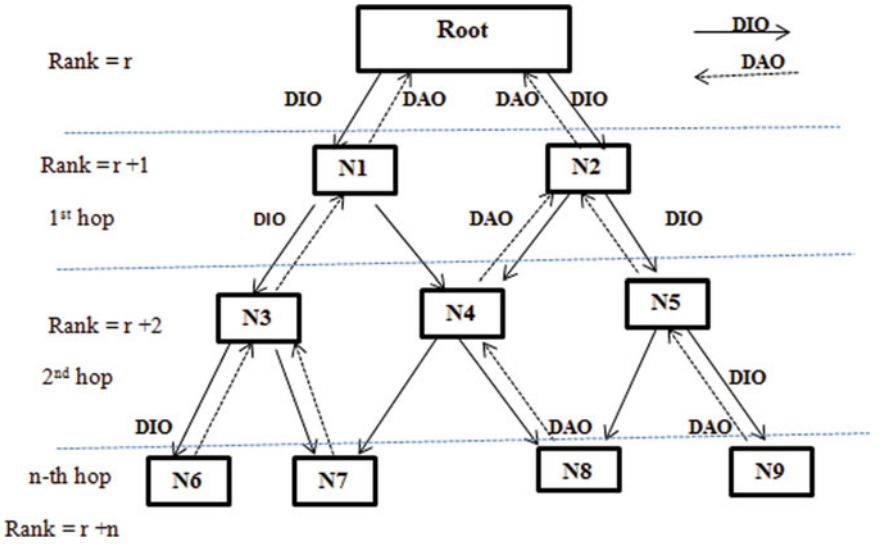


Fig. 78.1 Control messages in RPL protocol

is also utmost important to prevent most of the hazardous vulnerabilities and to safeguard the network. This in turn gives the significant improvements in the network performance.

In this paper, we introduce a decentralized blockchain-based RPL routing using smart contracts (smartRPL) which allows IoT network nodes from distinct vendors to trust one another and cooperate during the data communication.

Figure 78.1 shows the broadcast of RPL protocol control messages during the path establishment

78.1.1 Routing Protocol for Low Power and Lossy Networks

The process starts with building the DODAG root. The root acts as a gateway which sends the DIO messages to its neighboring nodes which has information like rank, RPL instance ID, version number, and DODAG ID; these DIO message are multicasted downwards depending on the received messages; the neighboring nodes decide if to join the root node or not. If the nodes want to join the DODAG root, then it has upward path, DAO messages are sent toward the root node, and at this point, the nodes compute their rank and select the parent based on the rank. The node selects the parent which has the lowest rank. Any new node can join the DODAG by sending the DIS message.

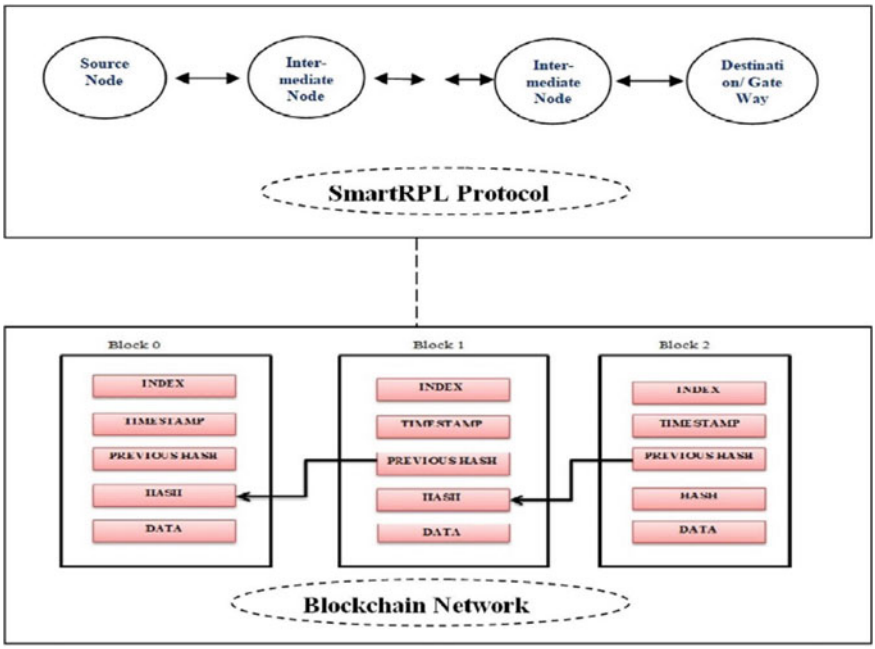


Fig. 78.2 Architecture of the proposed system

78.1.2 Blockchain Technology

Blockchain [3] was invented by Satoshi Nakamoto. It is chain of blocks which are interlinked to one another with the help of previous hash in the backward direction, every block contains details of the transactions, and this data which is stored in the blockchain network cannot be altered by a single node because of its distributed nature. Each block consists of index, timestamp, previous hash, hash, and data; there is unique ID for each and every block, and these are present in the sequential pattern. Figure 78.2 gives the structure of blockchain network with the number of elements that are present in each block in the network and shows the link among blocks.

78.1.3 Ethereum

Ethereum was discovered by Buterin [4]; it is an open source and decentralized platform that executes the programs called smart contracts. Ethereum is a platform specifically designed to build decentralized applications which provides transparency, availability, and auditability and removes intermediaries and allows users

to interact with each other directly through a global network. Ether is the main cryptocurrency which is used to pay transaction fee, and the average blocktime is 10–15 s. Ethereum can store transaction and documents.

78.1.4 Smart Contract

A smart contract is an agreement between two parties inside the smart contracts; there are set of conditions. All the contracts are based on terms, and conditions are recorded explicitly once the data written in the contract cannot be changed, updated, or deleted; they maintain the local copy in each block; it does not require any third party, and the data is in the encrypted form. Smart contracts are written in strength programming language [5] using a Remix IDE [6]. These smart contracts are deployed in ethereum platform, and to deploy the smart contracts, it requires the transaction fee. These are secure, trusted, and correct to store the data.

78.2 Problem Statement

IoT-LLNs are connected to various nodes in the network and exchange the data. These nodes should support parametric- constrained routing where nodes should select the best path to pass the data within the network. RPL routing protocol is used for LLNs, and it is vulnerable to different attacks [7] because of its resource-restraint nature to prevent these attacks; there are cryptographic algorithms, but these are managed using central authority, and these are difficult and expensive to implement. So, to support the trust between the nodes without any central authority, we propose the decentralized approach to find the best route to pass the data from source to destination, and it ends the single-point failure.

1. We propose a decentralized approach to set up a secure path from source to destination.
2. Performance metrics are considered to evaluate the performance.
3. Smart contracts are used to secure the data, and these are deployed in ethereum platform.

78.3 Literature Review

Few works that have been found in the literature focuses RPL security using various approaches. However, to the best of our knowledge, none of the work has been found focusing RPL routing security using blockchain-based smart contracts mechanism.

Dvir et al. proposed (VeRA) to protect the RPL protocol from the rank attack and version number attack; it prevents the attack node from changing version number and

rank; it is implemented using hash chain method [8], but this approach is vulnerable to rank attack.

Glissa et al. [9] proposed an approach to prevent rank modifications using threshold and hash chain authentication mechanism; it limits the rate of increase or decrease of rank values; it uses one-way hashed function such as secure hash algorithm 1 (SHA1). However, this approach causes the added overheads in the start to use hashing techniques.

Perrey et al. [10] proposed an approach TRAIL to solve the topology inconsistencies and discovers the fake nodes in the network. However, there are message overheads and increase the network size.

Ramezan et al. proposed a protocol [11] to prevent black hole and gray hole attacks using blockchain technology; it is used to find the trusted path between the source to destination. However, there are security issues in this model.

Conti et al. proposed Split [12] to secure the network from rank and Sybil attacks; a lightweight remote confirmation technique is used to prevent the network, but performance metrics are not considered in this approach.

Zhong et al. proposed an approach named Sprite [13] which is for mobile ad-hoc network. This approach needs the central authority to send the proof of message that data packet is received; it is vulnerable if the malicious node mounts the data to get there rewards.

78.4 Proposed System

In the proposed model, each IoT node creates the smart contracts to store the data; it has various parameters to find the trusted route and pass the data from source to destination. It uses RPL control messages to setup the path. It has different states to verify the route and pass the data. Figure 78.2 shows the structure of the proposed method which has a communication link (dotted lines) between the IoT network running RPL protocol to check the smart contracts present in the blockchain network for the correctness of the fields of control packets received via regular operations of routing protocol. If the fields of the control packets received by neighboring nodes are correct after checking with the information present in the smart contracts, then the intermediate node offers route via node through which it receives unmodified control packets. This approach continues till the secure path is offered from source to destination.

1. **Route Request**: To establish the secure route to the destination, we have to find the best path from source node to destination, and the source IoT node sends the DIO message which contains the information of contract address of the source node, rank, hop count, tokens, route request bond, and route request validity. As a bond, source IoT node transfers some of its tokens to the intermediate nodes to find the optimal route to destination and earn tokens. Route request bond also specifies the time until which the request is valid.

Table 78.1 Protocol parameters

S. no.	Parameter	Description
1	Contract address	Address of node
2	Rank	It calculates the individual node position from the root node
3	Hop count	It is count of hops from source node
4	State	Indicates the state
5	Route request bond	No. of tokens paid by source node to intermediate node as body
6	Route request expiry	It is route requested expiry time
7	Timestamp	Time at which the block is created in blockchain

2. **Route Offered**: The neighboring nodes send the valid route to the destination by DAO message and sends its information to its parent node maximum of three nodes which are stored in the smart contract. If the neighboring nodes are unaware of the routes and wants to join the network, it can pass the DIS message.
3. **Route Accepted**: The source IoT nodes verify the conditions to accept the route which is offered it selects based on its internal policies such as rank, hop count, and low-cost route from the offered routes to increase the throughput and security.
4. **Route Passed**: It selects the valid route based on the conditions, and the state of the contract is changed to data passed; if any node is unable to relay the packets, then the node's address is listed as blacklist address.
5. **Stop**: Anytime the source node can stop the routing process (Table 78.1).

78.5 Performance Evaluation

To evaluate smartRPL method, simulation study is performed using Cooja simulator with Contiki OS 2.6 [14] setup with 64 nodes in a grid structure and IEEE (Institute of Electronics and Electrical Engineers) 802.15.4 MAC (media access control). This implements the RPL routing protocol with an attack scenario and SRPL operations. Since our proposed work is blockchain based and do not allow the attackers to change the contents present in the RPL control packets, it ensures secure and reliable path establishment. This allow us to analyze the performance of the proposed method in comparison with the other two mentioned approaches. During the performance analysis study, it has been clearly seen that the proposed method has shown the significant improvements in terms of the key metrics such as packet delivery ratio (PDR) and throughput. The proposed model (smartRPL) has been evaluated and compared with RPL protocol with rank modification attack and SRPL [9] protocol. The results are significant and outperforms the existing two methodologies.

78.5.1 Packet Delivery Ratio (PDR)

Ensuring the secure path establishment among nodes in the network guarantee to offer higher packet delivery ratio. This can be defined as:

$$\text{Packet Delivery Ratio} = \frac{\text{No. of packets received}}{\text{No. of packets sent}}$$

As the number of attacker nodes increases, the proposed model shows higher packet delivery rate when it is compared with the RPL with rank attack and SRPL methodologies.

78.5.2 Throughput

Since the proposed mechanism offers blockchain-based secure routing approach, it do not allow the attacker nodes to change the control packets information. Thus, it can offer higher overall network throughput performance. Throughput can be defined as an average number of data packets received by destination per unit amount of time slot.

$$\text{Throughput} = \frac{\text{Data packets successfully received by destination}}{\text{Time taken}}$$

Simulation results show that the proposed model is highly effective in protecting the network against modification attacks as the blockchain has a public ledger system and undeniable properties. This reports better performance outcomes as shown in Figs. 78.3, 78.4, and 78.5. These simulations are performed in Contiki operating system [14].

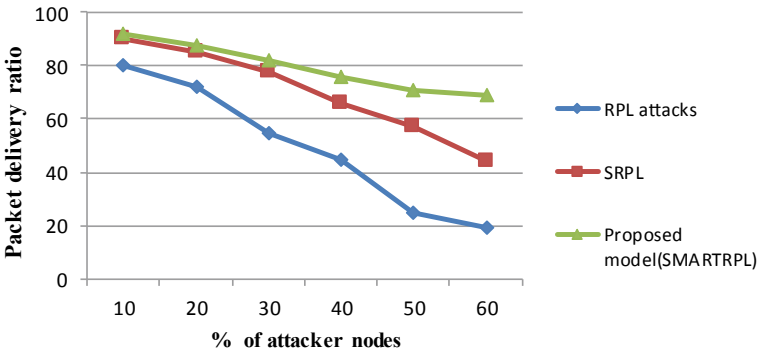


Fig. 78.3 Packet delivery ratio versus no. of attacker nodes

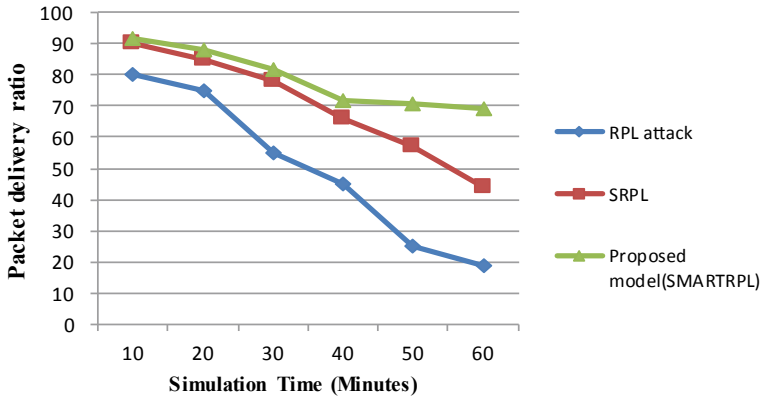


Fig. 78.4 Packet delivery ratio versus simulation time

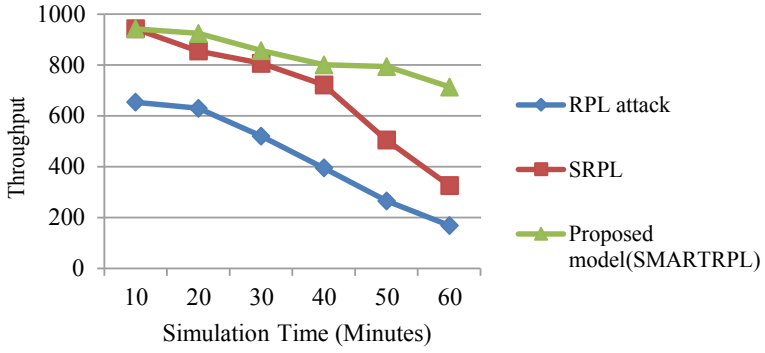


Fig. 78.5 Packet delivery ratio versus simulation time

78.6 Conclusion

RPL is the most promising routing protocol because of its efficiency for IoT-LLN devices describes how the nodes communicate with each other from source to destination. In the proposed system, a novel blockchain-based smart contracts method is used which is a decentralized, distributed ledger, and it creates a trustworthy IoT ecosystem which is capable to eradicate the risk of central authority and the nodes in the network; it establishes a secure route to pass the data from source IoT node to the destination node. This approach does not require a node to be equipped with a specialized hardware and avoid the RPL control packets from being changed by the attacker nodes during transmission without using more complex cryptographic algorithms. It uses smart contracts between node to node and checks for the correctness of control packets fields to set up a valid secure route between IoT nodes to gateway

or root and vice versa to pass the data securely during the data transmission phase in the network.

References

1. Airehrour, D., Gutierrez, J., Ray, S.K.: Secure routing for internet of things: a survey. J. Netw. Comput. Appl. **66**, 198–213 (2016)
2. Winter, T., Thubert, P., Brandt, A., Hui, J., Kelsey, R., Levis, P., Pister, K., Struik, R., Vasseur, J., Alexander, R.: RPL: IPv6 routing protocol for low-power and lossy networks (rfc 6550) (2012)
3. Zheng, Z., et al.: An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE International Congress on Big Data (Big Data congress). IEEE (2017)
4. Buterin, V.: Ethereum: A Next-generation Smart Contract and Decentralized Application Platform (2014). https://github.com/ethereum/wiki/wiki/%5BEnglish%5D-White-Paper
5. https://solidity.readthedocs.io/en/v0.4.24/introduction-to-smart-contracts.html
6. https://remix.ethereum.org/
7. Mayzaud, A., Badonnel, R., Chrisment, I.: A Taxonomy of Attacks in RPL-Based Internet of Things (2016)
8. Dvir A., Buttyan, L.: VeRA-version number and rank authentication in rpl. In: 2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems. IEEE (2011)
9. Glissa, G., Rachedi, A., Meddeb, A.: A secure routing protocol based on RPL for internet of things. In: 2016 IEEE Global Communications Conference (GLOBECOM). IEEE (2016)
10. Perrey, H., et al.: TRAIL: topology authentication in RPL(2013). arXiv:1312.0984
11. Ramezan, G., Leung, C.: A blockchain-based contractual routing protocol for the internet of things using smart contracts. Wirel. Commun. Mob. Comput. (2018)
12. Conti, M. et al.: SPLIT: a secure and scalable RPL routing protocol for internet of things. In: 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE (2018)
13. Zhong, S., Chen, J., Yang, Y.R.: Sprite: a simple, cheat-proof, credit-based system for mobile ad-hoc networks. In: IEEE INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428), vol. 3. IEEE (2003)
14. Get Started with Contiki. https://www.contiki-os.org/start.html

Chapter 79
Customized Music Classification and Recommendation System Based on Classifiers of Neural Networks and Sensor Embedded on Smart Devices

V. Roopa, A. ChristyJeba Malar, R. Rekanivetha, R. Thanga Pradeep Kumar, R. Sarveshwaran, and A. Prithiksha Parameshwari

Abstract Advancements in the process of embedding sensors in vehicles are helpful in recognizing the emotions of the person for annotating large music data set. The existing automated music playlist consists of collaborative filtering or content-based approaches in providing a fitting extension to a given playlist. This paper proposes a contextual-based music recommendation system that recommends songs based on the learning from user emotion. The user emotion is classified by the sensor embedded in the vehicle or the wearable computing device which is attached to the photo plethysmography (PPG) physiological sensors. This will recommend the songs to the user based on the context from the classified dataset. Prior to recommendation, songs are classified into genres considering the similarity in temporal structure and emotion classification is based on sensor values. This paper further proposes genre classification process with two steps comprising of feature extraction and music genre classification. The emotion classification from sensor values are may be based on decision tree, and Support vector and KNN algorithms of machine learning. Music

V. Roopa · A. ChristyJeba Malar (✉) · R. Rekanivetha · R. Thanga Pradeep Kumar · R. Sarveshwaran · A. Prithiksha Parameshwari
Department of Information Technology, Sri Krishna College of Technology, Coimbatore, Tamilnadu, India
e-mail: a.christyjebamalar@skct.edu.in

V. Roopa
e-mail: v.roopa@skct.edu.in

R. Rekanivetha
e-mail: rekharajkumar2000@gmail.com

R. Thanga Pradeep Kumar
e-mail: thangapradeep007@gmail.com

R. Sarveshwaran
e-mail: sarvesh.rajendran@gmail.com

A. Prithiksha Parameshwari
e-mail: prithiksha99@gmail.com

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_79

genre classification and recommendation with emotional and stress management provides a relaxation for the users while driving.

79.1 Introduction

Any song lasts for minutes, which is longer in duration. So, audio signal which is given as input for each and every song contains millions of samples with a sample frequency which is higher of about 10 khz. We prefer splitting a song into many frames so as to enable subsequent timbre extraction at frame level to signal analysis. This may lead to an efficient extraction process since we may require the application of spectral analysis only to signals of short time. Next, we have to choose an effective method for modeling the timbre features with the duration of 10–100 ms. The statistical properties of non-statistical series depend on time. Therefore on applying feature analysis at different frame level with the same audio signal. Thus timbre information is captured in a reliable manner. The wearable computing devices have sensors to measure heart rate, pulse rate and also provide the basis for physiological interpretations.

A feature content which is rhythmic includes the information about the exploration of beat tracking and tempo. Therefore beat spectrum is used for representation of rhythm. Frequency information is dealt by pitch content features obtained by using techniques of pitch detection. Nowadays deep learning techniques are explored regarding music genre classification. Thus the system provides the set of features for music which is direct and enables modeling. It explores the usage of Support vector in classification of music genre through K-Nearest Neighbour and Gaussian models. This music recommendation model extracts the correlation between the user emotion and the music genre.

Smart cars can have automatic play system enables such feature by extensively reading the sensor inputs and play the songs automatically with these predicted values. This customizes the genre of music being played and oozes out the imbalance which is the need of the hour. The proposed Machine learning algorithms emphasizes greater accuracy. The genre of music played in the proposed system used Convolution Neural Network (CNN) with classifiers as hidden convolution layers.

79.2 Literature Survey

Any song lasts for minutes, which is longer in duration. So, audio signal which is given as input for each and every song contains millions of samples with a sample frequency which is higher of about 10 khz. We prefer splitting a song into many frames so as to enable subsequent timbre extraction at frame level to signal analysis. This may lead to an efficient extraction process since we may require the application of spectral analysis only to signals of short time. Next we have to choose an effective method

for modeling the timbre features with the duration of 10–100 ms. The statistical properties of non-statistical series depends on time. Therefore on applying feature analysis at different frame level, we assume audio signals to be stationary. Thus timbre information is captured in a reliable manner. A feature content which is rhythmic includes the information about the exploration of beat tracking and tempo. Therefore beat spectrum is used for representation of rhythm. Frequency information is dealt by pitch content features obtained by using techniques of pitch detection. Not much work is explored regarding music genre classification. Thus the system provides the set of features for music which is direct and enables modeling. It explores their usage for classification of music genre through K-Nearest Neighbour and Gaussian models.

Fluctuation patterns from "focus" and "gravity". The combination of similarity and fluctuation patterns is mainly done for the addition of complementary information. Since it is sensitive to the effects that were produced, there is no usage of descriptors for overall loudness. There is no usage of spectral descriptors too as they are covered by the measure of spectral similarity.

A small audio segment is mainly used as it contains the information for characterizing the contents of a song because repetitions may be observed inherently in a musical structure. This idea is preferred as the computation required can be reduced. Proposed algorithms mostly use a small segment from audio for every title. To avoid introductions, 30s segment starts 30s after the piece is started.

Music can be categorized into genres in different ways. Music genre classification is useful to organize albums and songs which can be able to place music recommendations to the customers while driving or in downhearted. There are many genres of music classifications like pop, classical, and folk are available. There are different tones available for each genre. Around world 40 million songs are available in the digital library. It is time-consuming job to classify the music manually by listening the music. The person must have patience to listen the music and knowledge about genres. Some companies use music classification to place recommendation systems to the users (ex. Spotify, Soundcloud, Apple Music). Song recommendation system estimates the correlation between the users past listening history and songs.

Biofeedback-based music recommendation system [1] which consists of sensor-equipped headphones Septimu that records the heartbeat rate at the server. The recorded contextual information of the person was analyzed to suggest the dynamic music to maintain the target heart rate. A drowsiness detection system was used which records electroencephalogram (EEG) reading recorded by a sensor [2]. Artificial Neural Network (ANN), Support vector machine (SVM), and k-nearest neighbor (kNN) are integrated into the classification system to classify the drowsy state of the driver. Genetic algorithms also used to adjust weights in the classifiers [3]. Brainwaves classification and music recommendation are highly proved by experiments. Abdul et al. [4] generate implicit user rating for the music by calculating the correlation between user psychological data and music feature. Deep convolutional neural networks extract user behavior whereas weighted feature extraction (WFE) generates implicit user rating for music to get the correlation.

Different recommendation algorithms are evaluated using different types of measurement [5] which did not care about user emotion. An automatic stress reliving

system was designed for music listeners [6]. It has a wireless module with finger type sensor, which converts heartbeat rate to stress index level. Liu et al. [7] recommends a music playlist for the person which dynamically changes based on user heartbeat measurement. Yoon et al. [8] have developed a music recommendation system by feature selection, contextual information and past playlist history. Rosa et al. [9] implemented a music recommendation framework comprises of user profile with music preferences. The user's sentiment is analyzed by extracting the user's social media post. The framework recommends songs based on the metric.

Music recommendation system [10] involves the tasks of genre classification, song identification, feature extraction, and chord recognition. Shazam Entertainment is a music recommendation service available over internet, identifying artist and entire song receiving an audio sample (www.shazam.com). By Example (QBE) is a music search service available in Shazam. Many classification types are possible in music genre listed as Music Classification, Artist Identification, Mood Classification, and Music Annotation. Classifier learning is important to map the feature vector to the output label. Many audio features are considered in learning. Short-term, long-term features, semantic features are the considered categories in audio. Spectral Bandwidth, Amplitude Spectrum Envelop, Spectral Flux, Spectral centroid are some of the timbre features considered in music classification. Different timbres are possible by different sound sources of music. Markov et al. [11] have used Gaussian Process in music classification and emotion estimation which performed consistently better than support vector machine. Some of the features considered in music signal processing are,

Mel Frequency Cepstral Coefficients (MFCCs): Describe the shape of the spectral envelop used in speech recognition as well as audio similarity measures.

Line spectral Pairs (LSP): It is used to represent linear prediction coefficient to represent spectral envelop of a song in compressed form. It is the main parameter used in speech coding or synthesis.

Timbre (TMBR): Ability to distinguish different instruments. It is a group of four scalar features which include spectral centroid, spectral rolloff, spectral flux, and zero crossings.

Spectral Flatness Measure (SFM): SFM measures the differentiation between tone line and noise-like sounds.

Chromagram (CHR): Related to pitch is a complement of tone height. In music era, Two notes are separated one or more octaves and produce the same chorma.

79.3 Proposed System

The system classifies the genre of the songs in the repository of a person's music application through a predefined dataset and from that Convolution Neural Networks

(CNN) that helps in the classification of the music through the music features like MFCC, Mel's spectrum. GTZan and Million Songs Data set (MSD) are the mainly used datasets. GTZan is preferred as it has been widely used by many researchers. The Million Song Dataset (MSD) contains the meta data and derived features of the track such as artist, terms of the artist, similar artists, danceability, energy, duration, beats, tempo, loudness, and time signature and release year. There are seven genres such as blues, electronic, jazz, metal, rap, reggae, and rock are available with 1000 tracks in MSD. Mel frequency cepstral coefficients are mainly used in preprocessing technique without performing result. The features of classification include the percentage of training and test samples, the linear separation features for algorithms like SVM (Support Vector Machines), PCA (Principal Component Analysis), as in Fig. 79.1 independent target features, the sample overfitting percentage, overall accuracy, and performance measures. These results had training and modified testing data and the results are based on some pre-requisite labels used for classification. The enhanced networking model for classification with hidden layers represented the artificial neural networks. The proposed system has the classification part of fully connected layers that provides the classified recommendations as the output. In this module the features of a song will be selected and evaluated. Rhythm of the song has unique content features are movement of music signals over time, the rhythm and its regular interval, cycle of the beat, the rhythmic tempo and the lapse time signature. The rhythm structure is considered as feature set which is extracted from beat histogram. The melody and harmony information are also considered as features set and are extracted using pitch detection techniques. The music genres consists of the following features including:

- Tempo: The tempo of a piece of music measured in beats per minute which determines the speed of playing. In this experiment, the tempo of the entire song is taken into account. Any audio editor software can be used to calculate the tempo value of a song.

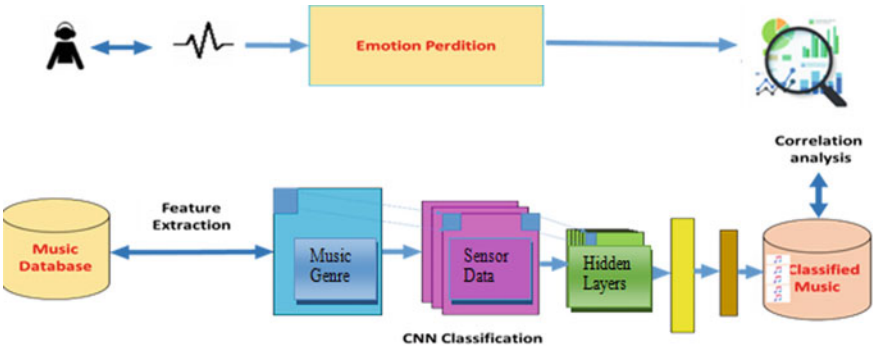


Fig. 79.1 Architecture diagram of proposed system using convolution neural networks and smart devices for music genre classification

- Beats: The basic unit of rhythm is beat. The beat value of the entire song if converted into a single digital value and is used as a feature for training the model.
- Chroma_feature: Chroma-based features represent the entire music in a condensed form. It is used in chord recognition, harmonic melody similarity analysis.
- RMSE: The Root Mean Square Error (RMSE) is used as an error metric. Residual measures the difference between the value predicted by model and the observed value.
- Spectral centroid: The spectral centroid is associated with brightness of the sound.
- Spectral Bandwidth: The spectral bandwidth refers to the spectral magnitude at one-half the peak maximum.
- Rolloff: Roll-off is the steepness of a transfer function with frequency.
- Zero_crossing_rate: The zero-crossing rate is the rate of sign-changes along a signal. This feature is highly used in MIR (Music Information Retrival).
- MFCC: Mel Frequency Cepstral Coefficients (MFCCs) The mel frequency cepstral coefficients (MFCCs) of a signal are a small set of features represents the short-term power spectrum of sound (usually about 10–20).

Sample music data set is fed into the classifier for training and it is classified according to the classifier results obtained. Once the classifier completes the classification, the sensor data from physical devices are fed to convolution layers for CNN algorithm implementation. The physical devices send the output value of simulation such as heart rate, pulse rate, physical activity, and stress rates at regular intervals to the trained neural network. These values are considered as the hidden layer of convolution pooling. The popular in store classifiers for songs and the emotion attached classified as in Fig. 79.2 shows the genre classification with sample music data set of 1000 songs. Techno—For party hard, for celebrating accomplishments. Listened usually with high volume. Euphoric Trance—To chill out after a long day or bad day at work. Mild volume with a sip of wine will be perfect. Uplifting music—Listened

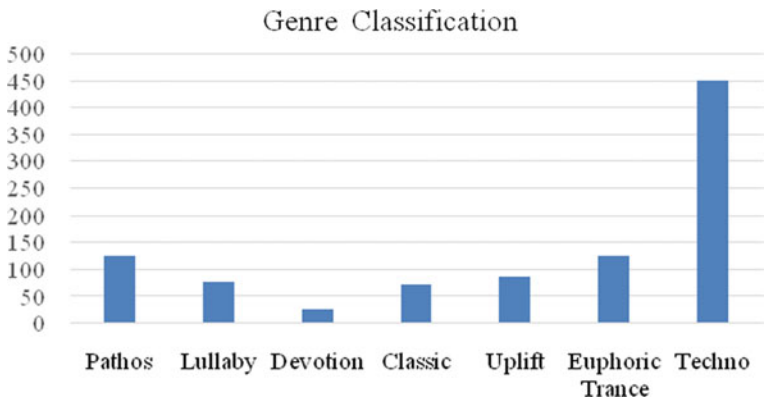


Fig. 79.2 Genre classification of sample data

during start of the day to derive needed energy to go after and accomplish the tasks planned for the day. Classical music—Music to be listened when in the mood of refreshing an orderly mind by enjoying the defined protocols in music. Western Classical and Carnatic music though they seem extreme they follow rigorous protocols which refreshes mind. Devotional music—This music is to direct mind and stay within the core belief system of the listener music. Lullaby—This music is for babies and toddlers. To make the child happy. They are very expressive and are considered impactful for child's growth. Folk music—The native music of the land. The rawness seen in this music is the beauty of this music. Brings out the natural characteristics of the listener in a more rustic way. It is also sung to ease the task at hand. Yelelo in India is sung by fishermen and farmers to ease their work and bring about bonding. Pathos—This is sad music. In other words, soothing a sad heart. This is music for the heart and not mind. Bereavement, broken heart, lost relationship are examples of pathos music situations.

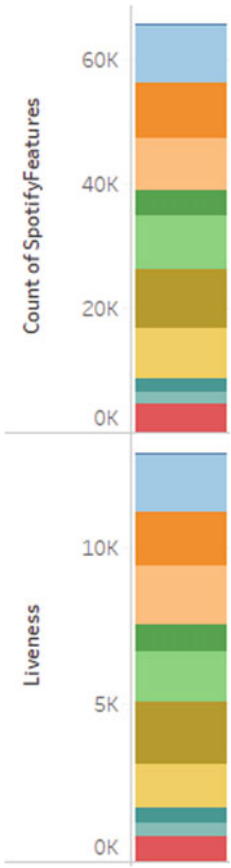


Fig. 79.3 Genre classification on liveliness from spotify features

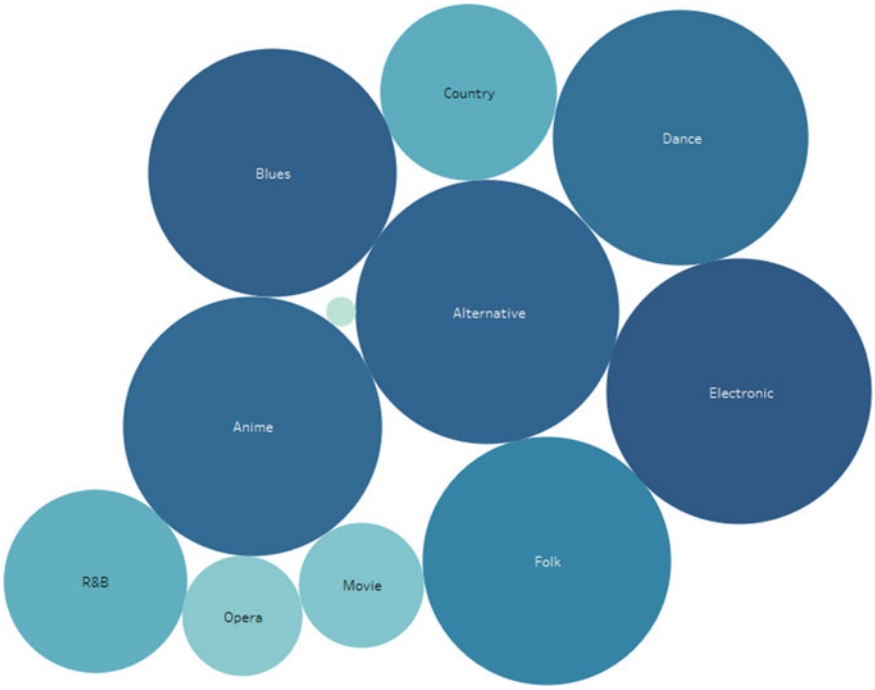


Fig. 79.4 Genre classification of spotify data

The existing model is implemented using Convolution Neural Network algorithm to find the genre of a given music with its feature data. CNN is used for classification [12]. Convolutional Neural Network is a type of deep learning algorithm consists of single input and single output layer with multiple hidden layers. Each hidden layer consists of series of convolution layer and followed by pooling layers. The commonly used activation function is RELU. The final convolution has back propagation to improve the final output. Mel Frequency Cepstral Coefficients (MFCC) features are extracted for each song. Fourier transform is applied to convert music into frequency. These extracted features are act as an input to neurons. Visual representation of frequencies of the music signal is represented by spectrogram. Table 79.1 shows the parameter for Training.

LibROSA is the python library needed to work on music data. Figure 79.5 shows the visualization of music spectrum.

79.4 Result and Discussion

The proposed system model with Convolution Neural Network (CNN) algorithm uses the feature and target values of the dataset are set as in Fig. 79.6, test and train

Table 79.1 Training parameters

Parameter	Value
Input layer	128 × 660 neurons
Convolutional layer	16 different 3 × 3 filters
Max pooling layer	2 × 4
Convolutional layer	32 different 3 × 3 filters
Max pooling layer	2 × 4
Convolutional layer	32 different 3 × 3 filters
Max pooling layer	2 × 4
Dense layer	64 neurons
Output layer	10 neurons for the 10 different genres

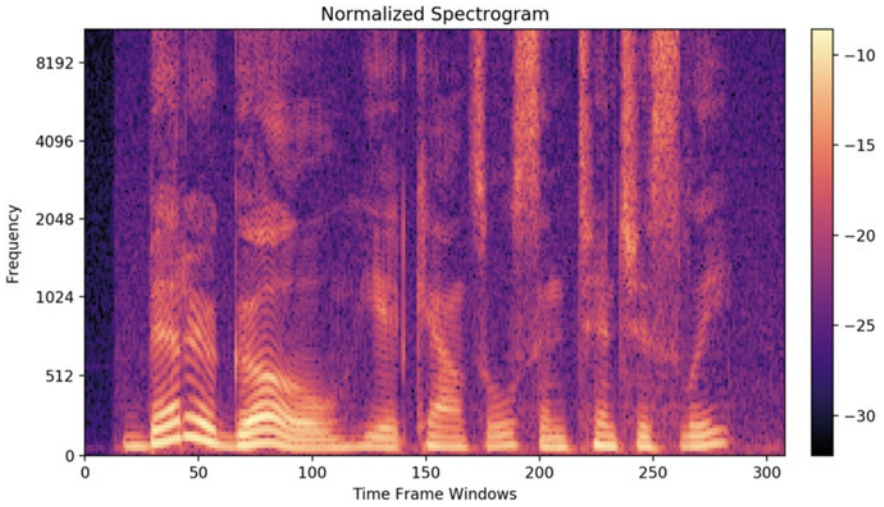


Fig. 79.5 Spectogram

data set are split, and in addition that a test size and random states that provides an optimal result.

The classified music data is further fed to these hidden layers. Thus an accuracy score is calculated. The music applications like JioSaavn, Wynk Music, Gaana, BigFM and many other applications have the online streamline of songs with more than 50 million songs in store. The accuracy and performance measures are the ease out of emotional stress and the involvement of music and mind to play together.

From the designed model the accuracy of the model obtained was 0.87 i.e., 87%. The input given to predict the genre was predicted successfully for all the inputs during the first trial and in the second trial for the second set of values the prediction for the 75% of the input data. The KNN with 75% accuracy was raised to the performance of 87% with CNN algorithm on convolution neural network as hidden

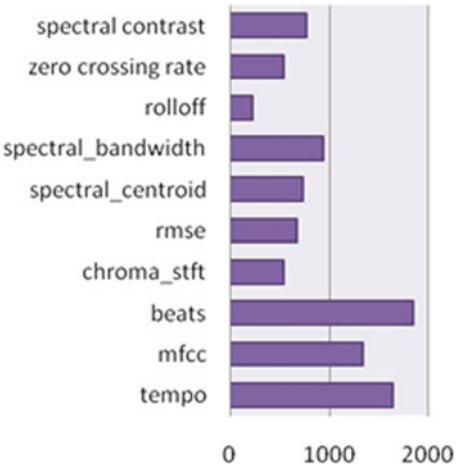


Fig. 79.6 Feature-based CNN Classifier

layer as in Fig. 79.7 has the accuracy performance measures. Table 79.2 shows the



Fig. 79.7 Performance measures

Table 79.2 Comparative study on different classifiers

Classifier	Accuracy
Decision tree	0.25
PCA	0.28
SVM	0.68
KNN	0.75
CNN	0.87

comparative study of classification algorithms is given.

79.5 Conclusion

This paper proposed a music genre classification and recommendation system using Convolution Neural networks. The feature sets are extracted using Mel Spectrum. The recent innovations in wearable devices are much needed for emotional predictions. Smart cars, smart homes, and many innovations are possible with the predictive values from the physical devices. The proposed system thus classifies music based on these sensor-enabled physical devices with music data set from many online stream applications. The system with features discussed above have many hidden layers as pooling over the convolution hidden layers. The classified music is thus played with the emotional simulation inputs across these hidden layers. The future scope is to build smart applications embedded with cognitive therapy for addressing emotional imbalance via music therapy. The increasing need of emotional intelligence with cognitive study can be developed using brain simulations over neural networks.

References

1. Nirjon, S., Dickerson, R.F., Li, Q., Asare, P., Stankovic, J.A., Hong, D., Zhao, F.: Musicalheart: a hearty way of listening to music. In: 10th ACM Conference on Embedded Network Sensor Systems, pp. 43–56. ACM (2012)
2. Liu, N.H., Hsieh, S.J., Tsai, C.F.: An intelligent music playlist generator based on the time parameter with artificial neural networks. Expert Syst. Appl. **37**(4), 2815–2825 (2010)
3. Liu, N.H., Chiang, C.Y., Hsu, H.M.: Improving driver alertness through music selection using a mobile EEG to detect brainwaves. Sensors **13**(7), 8199–8221 (2013)
4. Abdul, A., Chen, J., Liao, H.Y., Chang, S.H.: An emotion-aware personalized music recommendation system using a convolutional neural networks approach. Appl. Sci. **8**(7), 1103 (2018)
5. Isinkaye, F.O., Folajimi, Y.O., Ojokoh, B.A.: Recommendation systems: principles, methods and evaluation. Egypt. Inform. J. **16**(3), 261–273 (2015)
6. Shin, I.H., Cha, J., Cheon, G.W., Lee, C., Lee, S.Y., Yoon, H.J., Kim, H.C.: Automatic stress-relieving music recommendation system based on photoplethysmography-derived heart rate variability analysis. In: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6402–6405. IEEE (2014)
7. Liu, H., Hu, J., Rauterberg, M.: Music playlist recommendation based on user heartbeat and music preference. In: 2009 International Conference on Computer Technology and Development, pp. 545–549. IEEE (2009)
8. Yoon, K., Lee, J., Kim, M.U.: Music recommendation system using emotion triggering low-level features. IEEE Trans. Consum. Electron. **58**(2), 612–618 (2012)
9. Rosa, R.L., Rodriguez, D.Z., Bressan, G.: Music recommendation system based on user's sentiments extracted from social networks. IEEE Trans. Consum. Electron. **61**(3), 359–367 (2015)
10. Deger, A., Yusuf, Y., Kamasak, M.E.: Emotion based music recommendation system using wearable physiological sensors. IEEE Trans. Consumer Electron. **14**(8) (2018)

11. Markov, K., Matsui, T.: Music genre classification using Gaussian process models. In: 2013 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1–6. IEEE (2013)
12. Costa, Y.M., Oliveira, L.S., SillaJr, C.N.: An evaluation of convolutional neural networks for music classification using spectrograms. Appl. Soft Comput. **52**, 28–38 (2017)

Chapter 80
Localization of Optic Cup-Disc in Retinal Images Using Morphological Filters for Glaucoma Detection



Deepika Pal, Vikrant Bhateja, Archita Johri, and Babita Pal

Abstract Glaucoma is one of the main diseases causing permanent blindness. It is marked by an increase in Intraocular Pressure (IOP) which results in changing the shape of Optic disc and Optic cup. Segmentation and Localization of Optic Disc and Optic Cup for the automated detection and diagnosis of Glaucoma is necessary. Morphological Filters with the method of thresholding of the retinal fundus images have been adapted for segmentation of Region of Interest (ROI). The quality measurement of the segmentation of the ROI can be determined by the use of Similarity Coefficients. The resultant segmented binary images have been obtained with the value of Dice and Jaccard representing overlapping and similarity check between the segmented image and the ground truth. Cup to Disc Ratio has been calculated for the temporary analysis of Glaucoma.

80.1 Introduction

Glaucoma is the second largest cause of permanent blindness all over the world. Glaucoma patients or suspects develop symptoms like in any other ophthalmic disease. Hence it is difficult for the doctor to diagnose it at an early stage to start treatment at the earliest. Glaucoma patient's optic cup and disc start changing shape and size, eye of the suspect swells, redness is also developed, etc. In order to analyze Glaucoma in any patient some important measures have to be considered such as inheritance of the disease from his ancestors [1], any other history of ophthalmic diseases, cup and disc measurements, etc. There are many techniques adapted for the segmentation and localization of ROI. Empirical Wavelet Transform and B spline coefficient had been deployed by Maheshwari et al. [2]. EWT method dealt with decomposition of

D. Pal · V. Bhateja (✉) · A. Johri · B. Pal
Department of Electronics and Communication Engineering, Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPC), Faizabad, Road, Lucknow, U.P. 226028, India

Dr. APJ Abdul Kalam Technical University, Lucknow, U.P., India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_80

signals for both public and private databases. Three-Fold and Ten-Fold Cross Validation had been considered for it. The decomposed components obtained are arranged in an order of low to high-frequency ranges. Superpixel based segmentation by using Simple Linear Iterative Clustering (SLIC) algorithm had been used by Cheng et al. and Zhou et al. [3, 4]. SLIC approach dealt with tuning of optic disc boundary. Generation of super pixels for retinal fundus images had been considered for Glaucoma detection. Combination of Hough Transform, Polar Transform, and Morphological Operations had been used by Zahoor and Fraz [5]. Optic Disc localization had been performed by Circular Hough Transform (CHT), adaptive thresholding had been used for the segmentation of Optic Disc. CHT was used for the analysis of circular objects or components of the retinal image. Local image fitting model with oval-shaped constraint (LIFO) has been introduced by Gao [6]. LIFO had been used for localization of optic disc. Contour was formed using thresholding methods in combination with saliency map detection. Other OD segmentation for the COSFIRE filter was narrated by Guo et al. [7]. Vasculature selective and Disk selective COSFIRE filters were adopted for localization of OD in Glaucoma detection. Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm had been used for extraction of ROI by Sevastopolsky [8]. In order to combat for demerits of previous work on CAD of Glaucoma, an improved method of segmentation and localization of ROI based on combination of thresholding and morphological operation of retinal fundus images has been proposed in this paper. The rest of the paper is organized in the following sections: Sect. 2 discusses the proposed approach for segmentation and localization using morphological filter, thresholding method, ellipse fitting IQA, and CDR evaluation, Sect. 3 discusses the outcome of the proposed technique, Sect. 4 discusses the conclusion of the proposed work.

80.2 Proposefd Segmentation Methodology Using Morphological Filters and Thresholding

A. Morphology

Morphological filtering is the branch of science that deals with various operations related to different shapes and structures in any analysis. Morphology is tool for extracting various objects of different shape and size. Changing the size of ROI according to the need of our study or analysis [9]. It is applied on the images with a particular shape of structuring element which is desired at the output. Structuring element are different shapes matrices used for extraction of any object, component or ROI from any image or signal for its further processing. They are broadly classified as flat and non-flat structuring element. These includes structuring element of various shapes such as square, rectangle, disk and ball, etc. [10]. The selection criteria for the shape and size of structuring element depends on the need for analysis in which it is being used. Often used morphological operations include Erosion, Dilation

[11], Opening and Closing, etc. Morphology has various applications such as Image Processing, Signal Processing, etc. [12].

B. Thresholding and Ellipse Fitting Method

Thresholding is the method of extraction of significant information from the images automatically. Thresholding is the method of segmentation of various objects or parts of the image containing relevant information for further analysis [13]. In the proposed methodology Global Thresholding has been used. It is considered when any components or classes of image have different information. Global thresholding is considered when differences between two classes or parameters are distinct enough to be divided into foreground and background pixels in a binary image [14]. In this method, only one threshold is considered over the entire image to depict difference in parameters, etc. Global Thresholding is represented by Eq. 80.1 [15]:

$$g(x, y) = \left\{ \begin{array}{l} 0 \text{ if } f(x, y) < T \\ 1 \text{ if } f(x, y) \geq T \end{array} \right\} \tag{80.1}$$

Ellipse fitting algorithm is used for smoothening the ROI based on a least square fitting algorithm. It is based on the best-fit curve that has a minimal sum of deviations squared from a given data. It allows fitting the ellipse on a certain data points in a particular ROI [16].

C. Proposed Methodology for Segmentation and Localization

The proposed approach for OD and OC segmentation for the diagnosis of glaucoma using combination of Morphological filter. Simple operations of Morphology such as dilation and binary area filtering have been adapted for the segmentation module. Dilation is the morphological operation for adding the pixels at the region of interest. Binary area filtering is used for the extraction of the considered region or the region of interest by removing other unnecessary components of the image [17]. OD segmentation has been done on the enhanced red channel images. Standard Deviation has been calculated from the enhanced red channel images then thresholding has been applied to it. Binary area filtering has been used for removal of objects other than the region of interest i.e., OD. The segmented binary image of OD has been subjected to ellipse fitting for the smoothing of the disc. Finally, the segmented OD has been obtained from the retinal fundus images. OC segmentation has been carried out in the same way by computing its standard deviation of the green channel enhanced image. Then thresholding has been carried out then the resultant has been subjected to the binary area filtering. The filtered image is subjected to dilation by considering the ‘disk’ type structuring element of size 3 [17]. Ellipse Fitting has been applied to the binary image obtained. Finally, the segmented OC has been extracted. The extraction of ROI has been shown in Fig. 80.1.

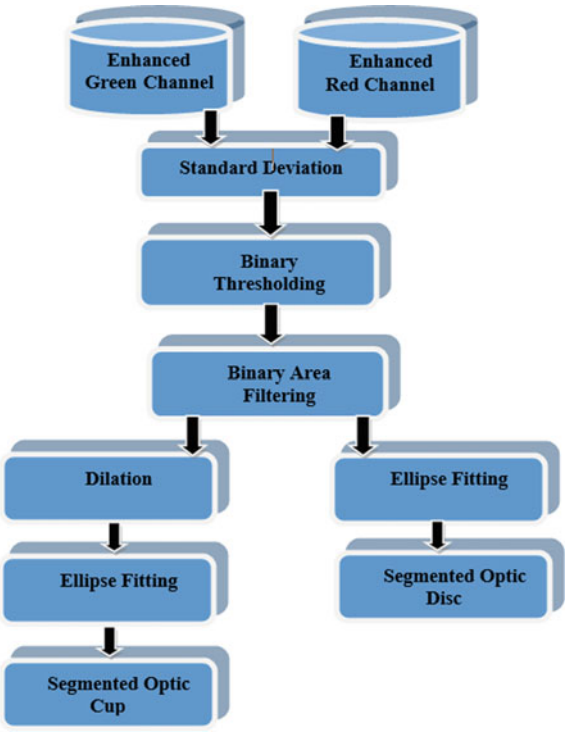


Fig. 80.1 Morphological segmentation module

D. **Proposed Evaluation of Cup to Disc Ratio**

The differentiation between glaucomatous and non-glaucomatous retinal fundus images has been done by computing Cup to Disc Ratio (CDR). CDR is calculated as the ratio of vertical diameter of cup to the vertical diameter of the disc. The value of CDR for retinal fundus image greater than 0.5 then the retinal fundus image is of glaucomatous eye else it is non-glaucomatous eye [18, 19]. The database considered for the CAD of Glaucoma is publicly available DRISHTI-GS. The dataset has been collected from Arvind Hospital, Madurai. It consists of total 101 images out of which 69 are Glaucomatous and 32 are normal retinal fundus images [20]. Figure 80.2 depicts calculation of OD and OC in a retinal fundus image.

The red circle in the image represents OD and green circle represents OC in the fundus image. For the computation of CDR vertical diameter of both disc and cup have been considered.

E. Image Quality Assessment using Similarity Coefficients

A similarity coefficient represents the similarity between two documents, two comparative quantities. A similarity coefficient is a function that computes the degree of similarity between a pair of text objects [21]. Similarity Coefficients for the assessment of segmentation have been proposed in this paper. In proposed method, Dice and



Fig. 80.2 OD and OC in the retinal fundus image [20]

Jaccard are the similarity coefficients considered for the assessment of the segmented ROI. The equations for Dice [22] and Jaccard [22] are as follows

$$\text{Dice} = 2\frac{A \cap G}{|A| + |G|} \tag{80.2}$$

$$\text{Jaccard} = \frac{|A \cap G|}{|A \cup G|} \tag{80.3}$$

80.3 Results and Discussions

The database consists of publicly available Drishti-GS dataset. It contains a total of 101 images. The dataset is divided into 50 training and 51 testing image sets. The images in the dataset are either glaucomatous or non-glaucomatous. All the images were considered for the analysis. Segmentation is the process of obtaining relevant information for the analysis of any problem. The segmentation results of the proposed methodology have been shown in Figs. 80.3 and 80.4. Figure 80.3 shows segmentation of optic cup whereas Fig. 80.4 shows segmentation of optic disc. Qualitative analysis of segmentation of ROI is performed using Sorenson similarity coefficients that are Dice and Jaccard. Segmentation results for Test_Image #A and Test_Image #B have been shown in Table 80.1.

IQA for the segmented ROI has been given in the Table 80.1. Various values of Dice and Jaccard represent that segmentation of the retinal fundus image has been carried out. Values of CDR depict that image is glaucomatous or non-glaucomatous.

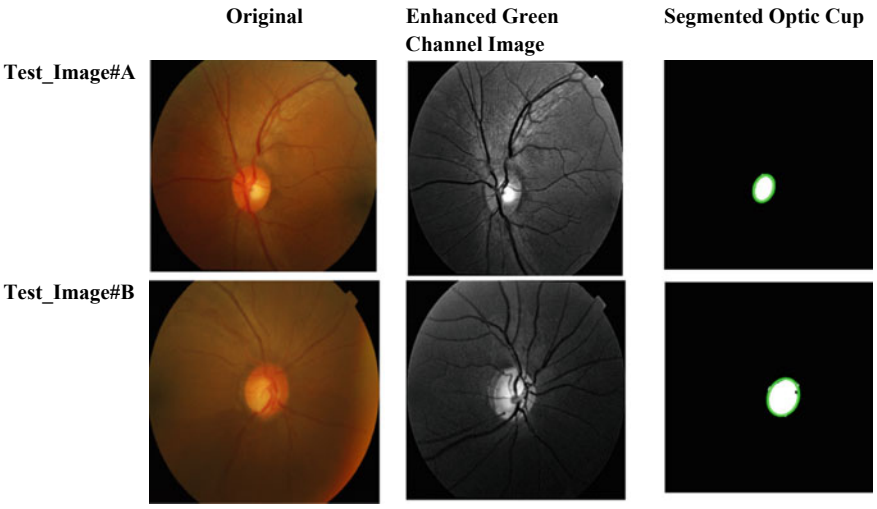


Fig. 80.3 Segmentation of optic cup

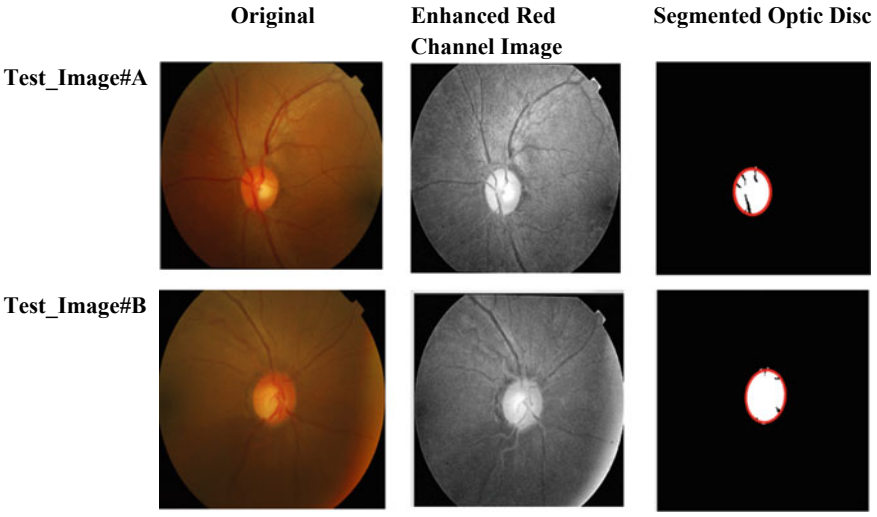


Fig. 80.4 Segmentation of optic disc

Table 80.1 IQA of Segmented ROI and CDR for glaucoma diagnosis

	Dice of OC	Jaccard of OC	Dice of OD	Jaccard of OD	CDR
Test_Image#A	0.0022	0.0011	0.0207	0.0104	0.6145
Test_Image#B	0.2793	0.1623	0.27292	0.15802	0.7594

80.4 Conclusion

CAD approach is a crucial step in diagnosis of various diseases thereby reducing the probability of error due to manual procedure or interference. The values of Dice and Jaccard are in the range between 0 and 1 indicating that segmentation has been done. Values of CDR is more than 0.5 in both the case indicating eye with less probability of affecting from glaucoma whereas the other one is glaucomatous with the value of CDR 0.7594. The proposed algorithm is simple for the segmentation of relevant objects or components of image. The proposed methodology can be improved by incorporating various features such as ISNT ratio, blood vessels removal, etc. Convolutional Neural Network (CNN) could also be used in order to increase the accuracy of the system. Various optimization algorithms could also be added to avoid tuning of the parameters.

References

1. Kavitha, S., Zebardast, N., Palaniswamy, K., Wojciechowski, R., Chan, E.S., Friedman, D.S., Venkatesh, R., Ramulu, P.Y.: Family history is a strong risk factor for prevalent angle closure in a south Indian population. Ophthalmology **121**(11), 2091–2097 (2014)
2. Maheshwari, S., Pachori, R.B., Acharya, U.R.: Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. IEEE J. Biomed. Health Inform. **21**(3), 803–813 (2016)
3. Cheng, J., Liu, J., Xu, Y., Yin, F., Kee Wong, D.W., Tan, N.-M., Tao, D., Cheng, C.Y., Aung, T., Wong, T.Y.: Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. IEEE Trans. Med. Imag. **10**(10), 1–15 (2013)
4. Zhou, W., Wu, C., Yi, Y., Duet, W.: Automatic detection of exudates in digital images using superpixel multi-feature classification. IEEE Access **5**(9), 17077–17088 (2017)
5. Zahoor, M.N., Fraz, M.M.: Fast optic disc segmentation in retina using polar transform. IEEE Access **5**(7), 12293–12300 (2017)
6. Gao, Y.: Automatic optic disc segmentation based on a modified local image fitting model with shape prior information. J. Healthc. Eng. Shenyang Liaoning China **3**, 1–10 (2019)
7. Guo, J., Azzopardi, G., Shi, C., Jansonius, N.M., PetkoV, N.: Automatic determination of vertical cup to disc ratio in retinal fundus images for glaucoma screening. IEEE Access **7**(1), 8527–8541 (2019)
8. Sevastopolsky, A.: Optic disc and cup segmentation methods for glaucoma detection with modification of U-net convolutional neural network. Pattern Recogn. Image Anal. Adv. Math. Theory Appl. 1–13 (2017)
9. Bhadauria, A.S., Nigam, M., Arya, A., Bhateja, V.: Morphological filtering based enhancement of MRI. In: Proceedings of 2nd International Conference on Computing, Communication and Control Technology (IC^4T), Lucknow, (U. P.), India, pp. 54–56, October, 2018
10. Structuring Element is available at: https://in.mathworks.com/help/images/structuring-elements.html
11. Bhateja, V., Devi, S.: A novel framework for edge detection of microcalcifications using a non-linear enhancement operator and morphological filter. In: IEEE 3rd International Conference on Electronics Computer Technology (ICECT), Kanyakumari, India, vol. 5, pp. 425–430, April, 2011

12. Srivastava, A., Alankrita, A., Raj, Bhateja, V.: Combination of wavelet transform and morphological filtering for enhancement of magnetic resonance images. In: International Conference on Digital Information Processing and Communications, Ostrava, Czech Republic, pp. 460–474, 7–9 July 2011
13. Verma, R., Mehrotra, R., Bhateja, V.: A new morphological filtering algorithm for preprocessing of electrocardiographic signals. In: Proceedings of 4th International Conference on Signal and Image Processing, pp. 193–201, January, 2013
14. Bhargavi, K., Jyothi, S.: A survey on threshold based segmentation technique in image processing. Int. J. Innov. Res. Devel. **3**(12), 234–239 (2014)
15. Castelman, K.R.: Digital Image Processing, Tsinghua Univ Press (2003)
16. Nayak, J., Acharya, R., Bhat, P.S., Shetty, N., Lim, T.C.: Automated diagnosis of glaucoma using digital fundus images. J. Med. Syst. **33**, 337–346, August, 2008
17. Gonzalez, R.C., Woods, R.E.: Digital image processing, 3rd ed. Pearson Education, Prentice Hall, Chap. 10, pp. 711–819 (2009)
18. Anusorn, C.B., Kongprawechnon, W., Kondo, T., Sintuwong, S., Tungpimolrut, K.: Image processing technique using cup to disc ratio. Thammasat Int. J. Sci. Technol. **18**(1), 22–34 (2013)
19. Singh, M., Virk, J.K., Singh, M.: A simple approach to cup-to-disk ratio determination for glaucoma screening, vol. 6, No. 2, pp. 177–182, September, 2015
20. Sivaswamy, J., Krishnadas, S.R., Joshi, G.D., Jain, M., Tabish, A.: DRISHTI-GS: retinal image dataset for optic nerve head (ONH) segmentation. In: IEEE 11th International Symposium on Biomedical Imaging, Beijing, China, pp. 1–4, July, 2014
21. Thada, V., Jaglan, V.: Comparison of Jaccard, Dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm. Int. J. Innov. Eng. Technol. (IJIET) **2**(4), 202–205 (2013)
22. Shrivastava, N., Bharti, J.: Empirical analysis of image segmentation techniques. In: International Conference on Smart Trends for Information Technology and Computer Communications, Jaipur, India, pp. 143–153, August, 2016

Chapter 81
Region Labeling Based Brain Tumor Segmentation from MR Images



Vikrant Bhateja, Mansi Nigam, and Anuj Singh Bhadauria

Abstract Brain tumor due to their increasing rate and high uncertainty has become a curse for mankind. For their effective diagnosis, automatic systems called Computed Aided Diagnosis (CAD) systems have developed that help in tumor analysis without manual interference. However, due to high variability in tumors, their segmentation from MR images is a challenging task. This paper proposes an improved tumor segmentation methodology that is an extension to simple thresholding technique. In this method, different regions of the binary images are labeled and are segregated on the basis of solidity and area. Then the region having solidity around 50% and maximum area is extracted as tumor. This segmented region is further dilated to include edema tissues surrounding the tumor. The performance of the methodology is justified by the results obtained in which only the tumorous region has been extracted indicating successful segmentation without inclusion of other brain tissues.

81.1 Introduction

Brain tumor is one of the most common causes for the increasing mortality among children and adults in the world. It is a group of abnormal cells that grows inside the brain creating pressure on the other tissues and the skull. With the increase in the number of people who develop brain tumor there is an immediate need to develop an automated system for early detection of brain tumor. Therefore, Computer Aided Diagnosis (CAD) systems are developed that are used to assist radiologists in interpreting the medical images with the help of automatic assisted tools. CAD technologies are employed to improve the diagnostic accuracy of the radiologists [1] by providing a second opinion about tumor diagnosis [2]. Magnetic Resonance Imaging (MRI) is one of the most reliable technique for brain tumor analysis because

V. Bhateja (✉) · M. Nigam · A. S. Bhadauria
Department of Electronics and Communication Engineering, Shri Ramswaroop Memorial Group of Professional Colleges (SRMGPC), Faizabad, Road, Lucknow, U.P. 226028, India

Dr. APJ Abdul Kalam Technical University, Lucknow, U.P., India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_81

it provides various images with different contrasting features thereby providing means for detailed analysis of tumorous tissues. One of the most crucial step of CAD systems is tumor segmentation from the brain image [3]. Segmentation is the process in which tumor region is separated from the entire brain image to detect its presence of an abnormal area. It is an important step and needs to be performed with utmost care because diagnosis at higher levels depends on the region extracted here. Incorrect segmentation may lead to false diagnosis of tumor which can be affected the patient severely [4, 5]. There have been a number of techniques use for tumor segmentation amongst which some of the existing approaches includes the concept of Hybrid and Variational Expectation Maximization, Tian et al. [6] utilized this technique to segregate the tumor region from other objects in the image. The proposed approach was easy to implement but did not consider the spatial relationship between neighboring pixels and was complicated to implement. Overcoming this problem Zhang et al. [7] presented a framework for medical image analysis which used Support Vector Machine (SVM) for tumor segmentation. Although this technique reduced time consumption but could not be used for inhomogeneous images which was generally a problem with medical images. To increase the reliability of the system Ji et al. [8] proposed fuzzy logic technique for segregation of infected tissues. Further, to extract the edge and texture information along with the infected region, Mohsen et al. [9] presented Feedback Pulse Coupled Neural Network based technique. However, CAD of brain tumor through MRI required more efficient and simple technique, abiding to which Somasundaram et al. [10] proposed a segmentation technique based on the adaptive intensity Thresholding. The above approach worked with dataset of any orientation and contrast but the selection of a particular threshold value to extract the exact tumor region excluding the brain tissues was a manual task and required clear knowledge about brain anatomy. Thresholding technique cannot be solely used to segregate the tumor region as it separates the infected area along with the objects of similar intensity. To brush-off the above challenges, this paper proposes a segmentation technique which employs thresholding and followed by region labeling techniques where the thresholded tissues are further subjected to another extraction on the basis of tumor solidity. This results into a segmented image containing only the abnormal tumor tissue. The remaining paper has been arranged into the following sections: Sect. 81.2 discusses the concept of region labeling and describes the proposed methodology; Sect. 81.3 explains and summarizes the results obtained and discusses them; Sect. 81.4 summarizes the conclusion of the work.

81.2 Proposed Methodology for Tumor Segmentation

81.2.1 Region Labelling

Image segmentation plays a vital role in medical image processing. It is the process of extraction of the region of interest by dividing the image into different regions

and the level up to which this division occurs depends on the problem being dealt. The segmentation is carried in order to extract the important details from which information can be perceived. For tumor diagnosis the segmentation is done to an extent that will yield the tumor area without the inclusion of brain tissues. The proposed segmentation methodology is based on Region Labeling. It is a technique in which different regions of a binary image are labeled so that different regions can be dealt separately and each region has its own identity. This helps in examining each region independently on the basis of certain properties and categorizing them. This distinction aids extraction of region that has properties resembling the tumor and therefore, tumor segmentation is done. In the proposed methodology, the property of tumor being exploited is solidity. Tumor solidity is 50% [11] which is very high compared to other brain tissues and hence region having this solidity level is identified as tumorous region using the process as illustrated in Fig. 81.1.

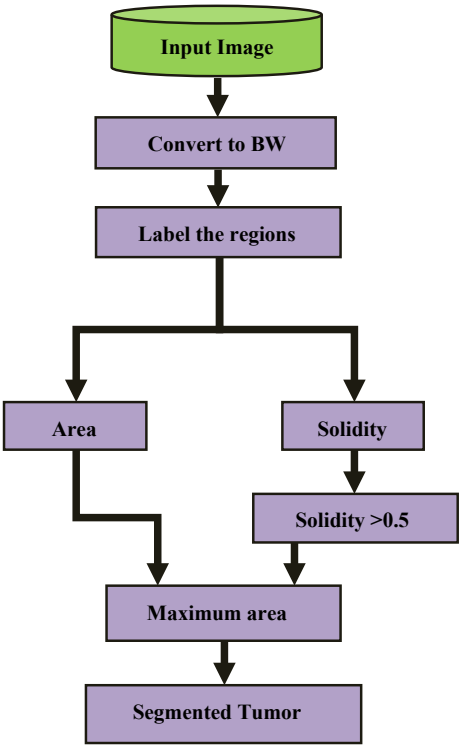


Fig. 81.1 Block diagram of proposed region labeling based segmentation technique

81.2.2 Proposed Region Labeling Based Tumor Segmentation

This paper suggests an improved method for segmentation of tumors from brain MR images using region labeling. In the first step, before labeling the regions, the image is converted into binary color map using thresholding which is a process of finding a specific intensity in a grayscale image that separates two distinct regions of an image, foreground, and background [12]. The proposed segmentation technique as shown in Algorithm 1 uses a threshold of value 0.3 for binarization because Region of Interest (ROI) i.e. Tumor is of very high intensity and is therefore always higher than this level. The pixel values greater than 0.3 are scaled to 1 which is white color (foreground) and those less than it are scaled down to 0 which is black color (background). Now the image consists of distinct regions of white color one amongst which is tumor region which is a dense region of uniform intensity and this image is subjected to region labeling. The white portion of the image is labeled into two classes, firstly on the basis of solidity and secondly on the basis of area. The regions having solidity greater than 50% are extracted and then amongst them the object of maximum area is obtained. This portion extracted is the segmented tumor region. In order to obtain the segregated tissues and balanced edges, the image is subjected to morphology-based operations. The binary image is dilated using a structuring element of size 5. Dilation [13] operation grows or thickens the objects in a binary image. It also bridges the gaps between the objects in the image [14] and is operated using Eq. 81.1. This process also incorporates edema, region of active cells surrounding the tumor region which otherwise will be removed due to its low-intensity value. The dilated image is complemented and the resultant image is subtracted from the input image which removes all the other tissues and retains the segmented tumor region.

$$X \oplus se = \{z|(se)_Z \cap X \neq \emptyset\} \tag{81.1}$$

Algorithm 1: Procedural Steps for Segmentation using Region Labelling.

BEGIN
Step 1: *Input* MR image [I].
Step 2: *Process* Convert [I] to Binary format [I_1] using 0.3 threshold.
Step 3: *Process* Label regions on the basis of area and solidity.
Step 4: *Process* Select the regions having solidity greater than 50%.
Step 5: *Compute* Area of selected regions.
Step 6: *Process* Select the maximum area region [I_2].
Step 7: *Define* Structuring element (*se*) of size 5.
Step 8: *Compute* Dilation [I_3] of image [I_2] using Eq. (1).
Step9: *Compute* Complement [I_4] of [I_3].
Step 10: *Output* Subtract Step 9 and Step 2.
END

81.3 Results and Discussions

81.3.1 Experimental Results

The brain MR images used for simulation of the proposed work are taken from the Whole Brain Atlas [15]. Segmentation of tumor region cannot be done only by thresholding because it extracts not only the tumor region but also other regions and tissues of similar intensity. To overcome this problem thresholding is followed by solidity-based segmentation. This method segregates the region of maximum area having solidity more than 50%. This leaves the resultant image with only a single region which is the region of interest i.e. tumor as shown in Fig. 81.2. The segmentation stage is the end of the brain tumor detection and first step in diagnosis. From the results, it is evident that the segmented region helps to identify whether the image is of tumor infected brain or not.

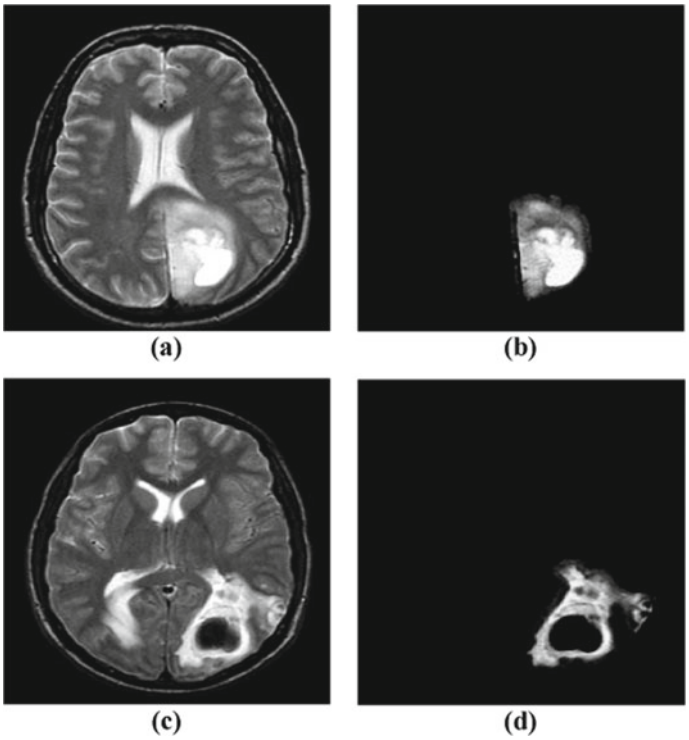


Fig. 81.2 **a** Original MRI image Test_Image#A. **b** Segmented Test_Image#A. **c** Original Test_Image#B. **d** Segmented Test_Image#B

81.3.2 Discussions

The above figure shows two MR images Test_Image#A and Test_Image#B and their respective segmented output images. The original Test_Image#A has a high, uniform intensity region at the lower half of the left occipital lobe and a slightly greyish edema region surrounding the tumor region. Along with this, there is a bright ventricular structure in the centre on the image which is a normal brain tissue. Upon segmentation the tumorous region is segregated from the other ventricular details with edema intact. In the Test_Image#B there are three main structures of similar intensities of which one is in the center, one at the lower half of the right occipital lobe and the other at lower half of the left occipital lobe. After segmentation, the homogeneous, dense region of maximum area present at the left occipital lobe along with the edema region is separated from the other objects. Therefore, the proposed methodology for tumor segmentation helps to obtain a separate tumor region but also preserve its original shape and structure.

81.4 Conclusion

A detailed analysis of the tumor region is necessary for proper detection and diagnosis through CAD systems which is possible by correct identification of tumor in the MR Images. In this paper, an extended approach of brain tumor segmentation from MR Images is suggested. Thresholding followed by region labeling is deployed which not only extract tumor solely by its intensity levels but also uses an important its property which is solidity that helps in further removal of similarly leveled brain tissues. Morphological dilation operator is further used that helps in inclusion of edema region surrounding the tumor that is of low intensity which can be observed in the results simulated and is usually omitted by thresholding. This inclusion is very important as it indicates the grade and exact dimensions of the tumor. Since in this technique precise details of tumor are also captured this is very helpful for diagnostic stage where features of tumor play a very important role to determine its grade.

References

1. Mohana, G., Subashini, M.M.: MRI based medical image analysis: survey on brain tumor grade classification. Biomed. Signal Process. Control **39**, 139–161 (2019)
2. Bhateja, V., Misra, M., Urooj, S.: Computer-aided analysis of mammograms. In: Non-Linear Filters for Mammogram Enhancement, pp. 21–27. Springer, Singapore (2020)
3. Basavaraju, H.T., et al.: Arbitrary oriented multilingual text detection and segmentation using level set and Gaussian mixture model. Evolut. Intell. 1–14 (2020)
4. Bhadauria, A.S., Bhateja, V., Nigam, M., Arya, A.: Skull stripping of brain MRI using mathematical morphology. In: Smart Intelligent Computing and Applications, pp. 775–780. Springer, Singapore (2020)

5. Bhateja, V., et al.: Two-stage multi-modal MR images fusion method based on parametric logarithmic image processing (PLIP) model. Pattern Recogn. Lett. (2020)
6. Tian, G., Xia, Y., Zhang, Y., Feng, D.: Hybrid genetic and variational expectation-maximization algorithm for Gaussian-mixture-model-based brain MR image segmentation. IEEE Trans. Inf. Technol. Biomed. **15**(3), 373–380 (2011)
7. Zhang, N., Ruan, S., Lebonvallet, S., Liao, Q., Zhu, Y.: Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation. Comput. Vis. Image Underst. **115**, 256–269 (2011)
8. Ji, Z., Suna, Q., Xiab, Y., Chena, Q., Xiaa, D., Feng, D.: Generalized rough fuzzy C-means algorithm for brain MR image segmentation. Comput. Methods Programs Biomed. **108**, 644–655 (2011)
9. Mohsen, H., El-Dahshan, E.A., Salem, A.M.: A machine learning technique for MRI brain images. In: Proceedings of 8th IEEE Conference on Informatics and Systems, pp. 161–16. Cairo, Egypt (2012)
10. Somasundaram, K., Kalaiselvi, T.: Automatic brain extraction methods for T1 magnetic resonance images using region labeling and morphological operations. Comput. Biol. Med. **41**(8), 716–725 (2011)
11. Radhi, A.A.: Efficient algorithm for the detection of a brain tumor from an MRI images. Int. J. Comput. Appl. **170**(10), 38–42 (2017)
12. Laddha, R.R., Ladhake, S.A.: A review on brain tumor detection using segmentation and threshold operations. Int. J. Comput. Sci. Inf. Technol. **5**(1), 607–611 (2014)
13. Arya, A., Bhateja, V., Nigam, M., Bhadauria, A.S.: Enhancement of brain MRT1/T2 images using mathematical morphology. In: Proceedings of 3rd International Conference on ICT, vol. 933, pp. 833–840. Springer Singapore (2019)
14. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, pp. 689–794. Pearson Education, Chap. 10 (2009)
15. The Whole Brain Atlas, https://www.med.harvard.edu/aanlib/home.html

Chapter 82
A Comparative Evaluation of Decomposition Methods Based on Pitch Estimation of Piano Notes



U. Vamsi Krishna, R. Priyamvada, G. Jyothish Lal, V. Sowmya, and K. P. Soman

Abstract Wavelet decomposition, variational mode decomposition, and dynamic mode decomposition are the latest signal processing tools that are recently being utilized in the music domain. Most of the work on these algorithms in music domain shows results based on pitch contour. None of the work mentions the effect of different frequency ranges (octaves) on these algorithms. In this paper, the evaluation is performed to understand the effect of different octaves on these decomposition algorithms based on pitch estimation of piano notes. Wavelet decomposition, variational mode decomposition, and dynamic mode decomposition methods are evaluated based on pitch estimation for different octaves. The purpose of this evaluation is to identify the most suitable method for pitch estimation. A comparative evaluation is performed on piano recordings taken from the database of the electronic music studio, University of Iowa. Absolute mean logarithmic error is used as the metric for evaluating the algorithms. The evaluation of algorithms is performed on seven different octaves ranging from 1 to 7. Variational mode decomposition performed better throughout the 7 octaves. Wavelet decomposition also performed well but was less accurate than variational mode decomposition. Dynamic mode decomposition was the least accurate among all the methods.

82.1 Introduction

In the music context, a note is a symbol denoting a musical sound. In western music, notes are ordered on a logarithmic scale [1]. The frequency of all the notes is standardized based 440 Hz which is the frequency of note A4. The frequency of note n is given as $440 \times 2^{\frac{n-49}{12}}$ Hz where n ranges from 1 to 88 on a standard piano. Pitch and frequency are related concepts, the pitch is considered to be the fundamental frequency of the note [2]. An octave is an interval between a musical pitch and another

U. Vamsi Krishna · R. Priyamvada · G. Jyothish Lal (✉) · V. Sowmya · K. P. Soman
Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
e-mail: g_jyothishlal@cb.amrita.edu

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0_82



Fig. 82.1 Piano notes with corresponding notations

pitch which is double its frequency [3]. There are 12 notes in an octave (denoted as C, C#, D...) as shown in Fig. 82.1 [1]. There are 7 complete octaves and each octave contains 12 notes. The 0th octave contains 3 notes that are A0, Bb0, B0, and the 8th octave contains 1 note, i.e., C8. More information regarding pitch and octaves can be found in [1]. The frequency of the piano notes increases exponentially, so the gap between the adjacent frequency of the note increases. For example, the frequency of A3 note 220 Hz which is in 3rd octave. The frequency next to A3 note, i.e., Bb3 is 233.1 Hz. The difference between the notes is 13.1 Hz. In contrast, for the 7th octave, A7 frequency 3520 Hz and the frequency of the next note, i.e., Bb7 3729 Hz. The frequency difference between the notes 209 Hz which is greater than in 3rd octave. If the +10 Hz error is added to both the notes, A3 frequency value will change 230 Hz which is close to the next note, i.e., Bb3. But for A7, if the same amount of error is added then it 2530 Hz which is still close to A7. This observation indicates that the effect of constant error decreases with respect to the octave. Typically used metric like RMSE might not be able to capture the error dependency. Logarithmic metrics can be used to deal with such dependencies. Section 82.4.1 talks about how the metric is selected for comparing the algorithms over octaves. There are many works based on pitch estimation algorithms. YIN is one of the classical pitch estimation algorithms [4] used for fundamental frequency estimation. The work presented in [5] shows how it fails for higher octaves of the piano. The parametric method in [5] also faces problems in detecting pitch for higher octaves. A genetic algorithm strategy is used to estimate the pitch for piano notes [6]. It is compared with three methods that are YIN, parametric F_0 estimation, and nonparametric F_0 estimation. Comparison results show that the genetic algorithm performs almost equal to the compared methods but parametric method performed better. Furthermore, pitch estimation is also required in pitch styling in speech signals [7]. Pitch styling is done to obtain a smooth pitch contour. Wavelet decomposition (WD) is utilized to estimate the pitches for pitch styling. Recent work on wavelet transform shows a modified Morlet wavelet filter to estimate the notes [8]. This work compares the proposed model with a fast Fourier transform and original Morlet wavelet for single pitch as well as multi-pitch estimation. The estimation of notes is done only for a few selected octaves. In one of the works, variational mode decomposition (VMD) is used to find the pitch contour in monophonic Turkish music [9]. The estimated frequencies are compared with the spectrogram of the signal. There is no metric evaluation mentioned for the estimated frequencies. Dynamic mode decomposition (DMD) is widely used for frequency estimation in many fields [10]. Recent work shows a windowed DMD approach which is a replacement for short-term Fourier transform to estimate the pitch contour [11]. As it is just a conceptual idea, there are no results that show concrete evidence of using the DMD algorithm for pitch estimation. A brief explanation of the three mentioned algorithms is given as follows.

82.1.1 Wavelet Decomposition

In wavelet decomposition, scaling and shifting are done to capture the time and frequency information present in the signal. There are many kinds of wavelets like Morlet, Haar, Daubechies, etc. [12]. The selection of these wavelets is dependent on the type of signal. Different level of decomposition leads to a different type of scaling. The generic criterion of scaling is $\frac{1}{2^n}$ where n is the level of decomposition. High scaling value will give higher frequencies accurately and lower scaling value for lower frequency (i.e., bigger window size will give details on lower frequencies and vice-versa).

82.1.2 Variational Mode Decomposition

VMD is a signal processing algorithm used to decompose an input signal into different band-limited sub-components which are called modes. These components are closely packed around the center frequency which is found in the decomposition process [13]. It uses optimization in frequency and time domain to find the optimal modes present in the signal. Here, the optimization is used to find the optimal functions. The mathematics of optimal function finding are known as variational calculus, hence, the name variational mode decomposition. It works on two constraints; the total sum of decomposed signals must be equal to the original and frequency variation from the central frequency (bandwidth) is minimum. The analytical form is obtained by using the Hilbert transform. Frequency mixing is done on the analytical function to find the central frequency present in the corresponding mode. Then, frequency spectrum shifting is done to the baseband by mixing complex exponential which is adjusted with the early estimated central frequency.

Mathematically, we can express as,

$$\min_{m_k(t),\omega_k} \sum_k \left\| \frac{\partial}{\partial t} \left[\left(\left(\delta(t) + \frac{j}{\pi t} \right) * m_k(t) \right) e^{-j\omega_k t} \right] \right\|_2^2 \tag{82.1}$$

$$\text{s.t.} \sum_k m_k(t) = f(t)$$

where $m_k(t)$ is the kth mode of original signal $m(t)$ with center frequency w_k and k is the integer index of the mode. Augmented Lagrangian multiplier is used to convert the constrained to unconstrained problem.

$$L(m_k, w_k, \lambda) = \alpha \sum_k \left\| \frac{\partial}{\partial t} \left[\left(\left(\partial(t) + \frac{j}{\pi t} \right) * m_k(t) \right) e^{-j\psi_k t} \right] \right\|_2^2 + \left\| f - \sum_k m_k \right\|_2^2 + \langle x, f - \sum_k m_k \rangle \quad (82.2)$$

where f is the original signal.

82.1.3 Dynamic Mode Decomposition

DMD is a data-driven method used to predict dynamics of the stationary system. It uses singular value decomposition (SVD) as its core algorithm and is used to estimate the low-rank dynamics [14]. Data snapshots from the signal are stored in the matrices X_1 and X_2. Where X_1 is the matrix containing 1 to n snapshots and the X_2 is the delayed version containing snapshots from 2 to $n + 1$.

$$X_1 = \begin{bmatrix} : & : & & : \\ x_1 & x_2 & \dots & x_n \\ : & : & & : \end{bmatrix}$$

And a delayed version of the snapshot is taken

$$X_2 = \begin{bmatrix} : & : & & : \\ x_2 & x_3 & \dots & x_{n+1} \\ : & : & & : \end{bmatrix}$$

The core part of DMD is to find a matrix A such a way that

$$X_2 \approx A X_1 \quad (82.3)$$

The eigenvalues of matrix A are used to estimate the frequencies f present in the signal.

$$f = \frac{imag(\log(\lambda))}{2\pi \Delta t} \quad (82.4)$$

where λ is the eigenvalues of the A matrix which is computed using SVD. Δt is the time-step of the signal. Further, details are given in the [15].

82.2 Dataset Description

Dataset is taken from the database of electronic music studios, the University of Iowa [16]. It contains .aiff stereo (left and right information) format 88 audio files. Each file contains an individual note of Steinway & Sons Model B piano and is recorded using Neumann KM 84 microphone. The dataset is noise-free. The sampling and bit rates of audio files are 44.1 KHz and 16-bit, respectively.

82.3 Methodology

All audio files in the dataset contained more silence regions after and before the original piano note, as shown in Fig. 82.2a. Different lengths of the window are applied on the signal to find out the region where the energy of the signal is high as shown in Fig. 82.2b. After finding the appropriate region, decomposition algorithms are applied. After analyzing the signal, the frequency with the highest energy is taken as the fundamental frequency. Figure 82.3 shows the flow diagram of the method.

82.4 Results and Discussions

82.4.1 Selection of Metrics

Root mean square logarithmic error (RMSLE), RMSE and mean absolute logarithmic error (MALE) are compared to show which metrics can capture the error dependency mentioned in Sect. 82.1. Equations 82.5, 82.6 and 82.7 are RMSLE, RMSE and MALE, respectively.

$$\sqrt{\sum \frac{(\log(y_{\text{ref}}+1)-\log(y_{\text{pred}}+1))^2}{N}} \tag{82.5}$$

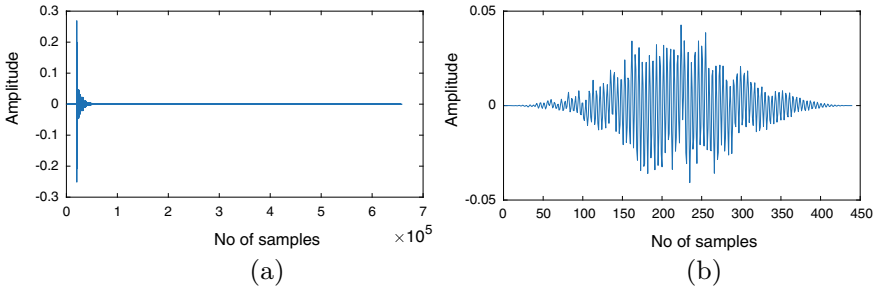


Fig. 82.2 **a** Original signal and **b** windowed signal

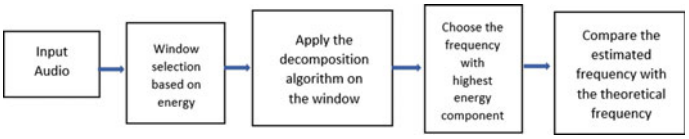
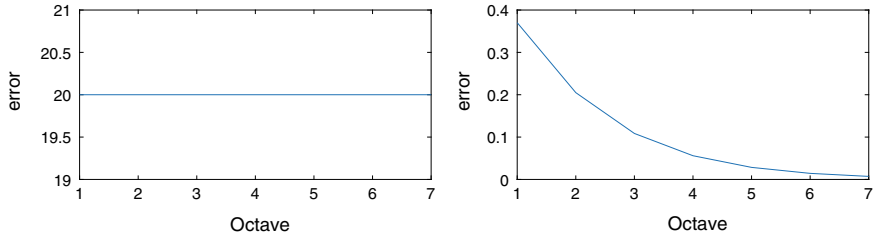
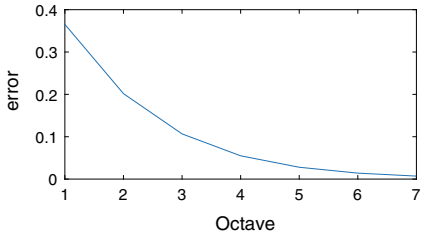


Fig. 82.3 Methodology



(a) RMSE (b) RMSLE



(c) MALE

Fig. 82.4 Comparison of metrics over 7 octaves

$$\sqrt{\sum \frac{(y_{\text{ref}} - y_{\text{pred}})^2}{N}} \tag{82.6}$$

$$\sum \frac{abs(\log(y_{\text{ref}}) - \log(y_{\text{pred}}))}{N} \tag{82.7}$$

where y_{ref} is reference value, y_{pred} is predicted value and N total number of values. A constant error of ± 20 Hz is added to all the 88 piano frequencies. As shown in Fig. 82.4a, RMSE remains constant throughout the 8 octaves. On the other hand, as shown in Fig. 82.4b, c, RMSLE and MALE are decreasing with respect to the octaves, respectively. The mathematical reason for RMSLE and MALE to work better is due the logarithm function utilized in both the metrics. The expression for the frequency of notes is given as $f(n) = 440 \times 2^{\frac{n-49}{12}}$, where n is the integer index of piano note. By applying logarithmic function on $f(n)$, we get

$$\log(f(n)) = \log(440) + \frac{n-49}{12} log(2)$$

where $\log(f(n))$ is linear in frequency and the difference between frequencies becomes equal for all octaves. MALE and RMSLE perform identically, but RMSLE uses square root, mean and square operations. On the other hand, MALE uses mean and logarithmic difference. This makes MALE computationally effective.

Table 82.1 Tunable parameters for WD

Octave	Fs (Hz)	Window length (s)	Decomposition level
1	441	0.5	1
2	4410	0.02	2
3	4410	0.5	3
4	4410	0.25	2
5	4410	2	1
6	22,050	0.4	2
7	44,100	0.02	2

Table 82.2 Tunable parameters for VMD

Octave	Fs (Hz)	Window length(s)	K	α
1	220	1	4	20,000
2	2205	0.2	1	15,000
3	2940	0.15	1	10,000
4	4410	0.1	1	10,000
5	4410	0.1	3	10,000
6	8820	0.05	5	10,000
7	8820	0.05	6	5000

Table 82.3 Tunable parameters for DMD

Octave	Fs (Hz)	Window length(s)	Data snapshot length (N)	D
1	221	4.9	100	1
2	294	3.74	100	1
3	441	2.49	180	1
4	882	1.13	180	1
5	2940	1.5	180	1
6	8820	0.12	100	2
7	44,100	0.02	180	1

82.4.2 Parameter Tuning

Tables 82.1, 82.2 and 82.3 show the tunable parameters for WD, VMD and DMD algorithms, respectively. Each algorithm is tuned based on the least MALE. Hanning, Hamming and Blackman windows are used but there is no significant change in the MALE. Applying any of the mentioned window will produce almost the same result. In this paper, Blackman window is used. Sampling rate F_s and window length are the common parameters for all the three algorithms. Table 82.1 contains the parameters

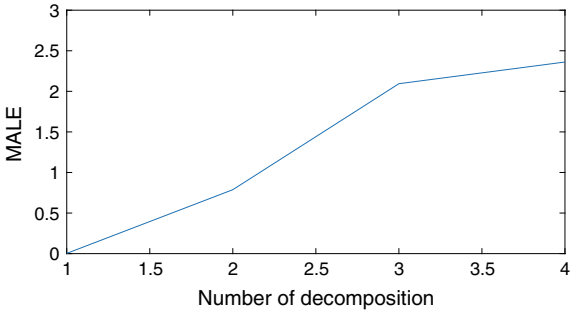
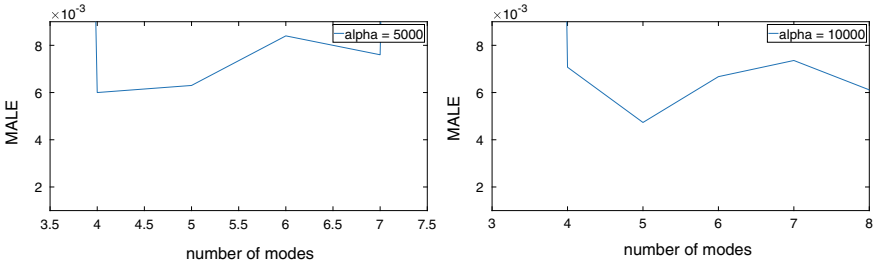


Fig. 82.5 WD parameter tuning for 5th octave

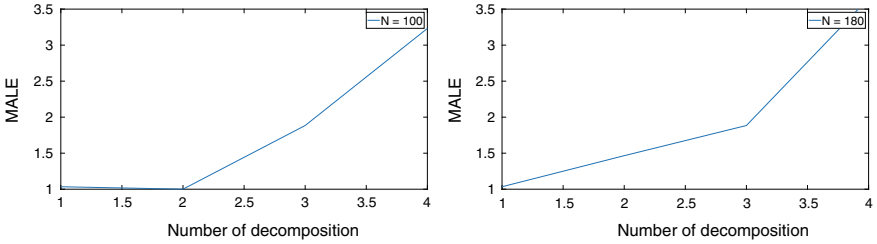


(a) K=4 and $\alpha = 5000$ produces least MALE of 0.006Hz

(b) K=5 and α=10000 produces least MALE of 0.0045Hz

Fig. 82.6 VMD parameter tuning for 6th octave

selected for the WD algorithm. Figure 82.5 shows the selection of the decomposition level with respect to the MALE for octave 5. First level decomposition is selected because it provides the least 0.003 Hz MALE, as shown in Fig. 82.5. Same criterion is used for all the other octaves. Second level decomposition is applied on 2nd, 4th, 6th and 7th octave. First level decomposition is applied on the 1st and 5th octave. The third level of decomposition is applied on the 3rd octave. VMD parameters are given in Table 82.2. K is the number of modes, and α is the bandwidth parameter. Figure 82.6 shows the tuning of parameters based on MALE for octave 6. Figure 82.6(a) shows that $K = 4$ produces the least 0.006Hz MALE for $\alpha = 5000$. Figure 82.6(b) shows that $k = 5$ and $\alpha = 10{,}000$ gives 0.0045Hz of MALE which is lower than the previous scenario. $K = 5$ and $\alpha = 10{,}000$ are selected for octave 6. Similarly, for all the octaves, same criterion is applied. α is selected as 5000 for the 7th octave. For 3rd, 4th, 5th and 6th octave, α is 10000 and for 1st and 2nd octave, α is taken as 20000 and 15000, respectively. DMD parameters are given in Table 82.3. Here, D is the number of modes and N is the length of snapshots. Figure 82.7 shows the parameters tuned for octave 6 based on MALE with snapshots 100 and 180 for comparison. For $D = 1$ and $N = 100$, DMD gives 1.005 Hz MALE, as shown in Fig. 82.7a. For $D = 1$ and $N = 180$ DMD gives 1.0065 Hz MALE, as shown in Fig. 82.7b. Comparing both the error,



(a) D=1 and N = 100 produces least MALE of 1.005Hz

(b) D=1 and N=180 produces least MALE of 1.0065Hz

Fig. 82.7 DMD parameter tuning for 3th octave

DMD performs better for $D = 2$ and $N = 100$ for 6th octave. Similarly, parameter for all the octaves is selected. For octave 6, the second decomposition is selected, and for the rest of the octaves, the first decomposition is selected. For octave 3rd, 4th, 5th and 7th, the N is taken as 180, and for octave 1st, 2nd, 6th, it is taken as 100. Table 82.4 shows that VMD performed well for all the octaves and overall accuracy is also high. The expression for accuracy is given as

$$\text{Accuracy} = \frac{N_{\text{est}}}{N_{\text{total}}} \times 100$$

where N_{est} is number of notes correctly estimated and N_{total} is total number of notes. As shown in Table 82.2, the bandwidth parameter α keeps on decreasing with an increase in octave. The higher the frequency, the lower the α value. The error for lower octaves is very high, this is because VMD cannot detect frequencies which are close to each other. From octave 1st to 4th, WD did not perform well with respect to VMD. From 5th to 7th octave, WD is able to estimate all the notes correctly.

82.5 Conclusions and Future Scope

The evaluation of all three algorithms shows that VMD performed better for all the octaves. WD also performed well but not accurate as VMD. DMD did not perform well for any of the octave, which makes DMD the least choice for pitch estimation application. While decomposing the signal, VMD has a frequency mixing criteria which can capture the fundamental frequency in conditions where harmonics are as strong as fundamental frequency. WD uses high pass and low pass filters for decomposition which can capture fundamental frequency given the harmonics are not too close. Unlike VMD and WD, DMD does not have any such property with respect to frequency filtering. This could be one of the reasons why DMD was not able to perform for any octave. The evaluation also shows how the algorithms are dependent

Table 82.4 Evaluation of the algorithms

Octave	VMD		WD		DMD	
	MALE	Correct notes	MALE	Correct notes	MALE	Correct Notes
1	0.3536	2	0.385	1	1.66	2
2	0.061	11	0.396	2	1.57	1
3	0.046	11	0.24	5	1.66	1
4	0.0043	12	0.23	8	1.21	3
5	0.0033	12	0.003	12	1.0003	2
6	0.0045	12	0.006	12	1.0005	8
7	0.0133	12	0.015	12	0.74	3
Total notes	72		52		20	
Accuracy (%)	85.71		61.9		23.81	

on sampling frequency and window length for different octaves. Further, experiments can be conducted to identify the optimal sampling frequency and window length to accurately estimate pitches for all the octaves. Multi-pitch estimation can also be evaluated on these algorithms based on MALE. Furthermore, noise can be added to evaluate the algorithms in different conditions. VMD's performance is not expected to reduce due to its wiener filter properties [13]. DMD is very sensitive to noise which can further lower its performance [17]. Evaluating the modified versions of these algorithms like multi-resolution DMD [18], elastic regression VMD [9], etc., can give better insights.

References

1. Klapuri, A., Davy, M.: Signal Processing Methods for Music Transcription. Springer Science & Business Media, Berlin (2007)
2. Goto, M.: A real-time music-scene-description system: predominant-F0 estimation for detecting melody and bass lines in real-world audio signals. Speech Commun. **43**, 11 (2004). https://doi.org/10.1016/j.specom.2004.07.001
3. Smith, W., Cheetham, S.: A Dictionary of Christian Antiquities. John Murray, London (1875). Archived from the original on 30 Apr 2016
4. Cheveigné, A., Kawahara, H.: YIN, a fundamental frequency estimator for speech and music. J. Acoust. Soc. Am. **111**, 1917–30 (2002). https://doi.org/10.1121/1.1458024
5. Emiya, V., David, B., Badeau, R.: A parametric method for pitch estimation of piano tones. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, 1. I-249 (2007). https://doi.org/10.1109/ICASSP.2007.366663
6. Inácio, T., Miragaia, R., Reis, G., Grilo, C., Fernandéz, F.: Cartesian genetic programming applied to pitch estimation of piano notes. In: IEEE Symposium series on computacional inteligence (SSCI) (2016)
7. Wang, D., Narayanan, S.: Piecewise linear stylization of pitch via wavelet analysis, pp. 3277–3280 (2005)

8. Kumar, N., Kumar, R.: Wavelet transform-based multipitch estimation in polyphonic music. Heliyon **6**(1) (2020)
9. Öztürk, B., Akan, A., Bozkurt, B.: Fundamental frequency estimation for monophonical Turkish music by using VMD. In: 2015 23rd Signal Processing and Communications Applications Conference, SIU 2015—Proceedings, pp. 1022–1025 (2015). https://doi.org/10.1109/SIU.2015.7130006
10. Mohan, N., Soman, K.P., Sachin, K.S.: A data-driven approach for estimating power system frequency and amplitude using dynamic mode decomposition. In: International Conference and Utility Exhibition on Green Energy for Sustainable Development (ICUE), vol. 2018, pp. 1–9. Phuket, Thailand (2018). https://doi.org/10.23919/ICUE-GESD.2018.8635792
11. Pogorelyuk, L., Rowley, C.: Melody Extraction For Mirex 2016 Using Dynamic Mode Decomposition. Mirix **2017**, (March 2016)
12. Soman, K.P., Ramachandran, K.I.: Insight Into Wavelets: From Theory to Practice. PHI Learning Pvt. Ltd. (2010)
13. Dragomiretskiy, K., Zosso, D.: Variational mode decomposition. IEEE Trans. Sign. Process. **62**(3), 531–544 (2014)
14. Schmid, P.J.: Dynamic mode decomposition of numerical and experimental data. J. Fluid Mech. **656** (2008). https://doi.org/10.1017/S0022112010001217
15. Kutz, J.N., et al.: Dynamic mode decomposition: data-driven modeling of complex systems. Soc. Ind. Appl. Math. (2017)
16. Mazunik, E., Cash, M.: University of Iowa Electronic Music Studios piano dataset. Available at http://theremin.music.uiowa.edu/MISpiano.html (2001)
17. Hemati, M.S., Rowley, C.W., Deem, E.A., et al.: De-biasing the dynamic mode decomposition for applied Koopman spectral analysis of noisy datasets. Theor. Comput. Fluid Dyn. **31**, 349–368 (2017). https://doi.org/10.1007/s00162-017-0432-2
18. Kutz, J., Fu, X., Brunton, S., Erichson, N.: Multi-resolution dynamic mode decomposition for foreground/background separation and object tracking, pp. 921–929 (2015). https://doi.org/10.1109/ICCVW.2015.122
19. Salamon, J., Gomez, E.: Melody extraction from polyphonic music signals using pitch contour characteristics. IEEE Trans. Audio Speech Language Process. **20**(6), 1759–1770 (2012)
20. Lal, G.J., Gopalakrishnan, E.A., Govind, D.: Epoch estimation from emotional speech signals using variational mode decomposition. Circuits Syst. Sign. Process. **37**, 3245–3274 (2018)

Author Index

A
Abdulla, Wissam Salih, 387
Achha, Nethaji, 25
Aditya Nag, M. V., 47
Akhilesh, D., 35
Akuli, Amitava, 573
Althaf Hussain Basha, Sk., 47
Amanchi, Uma Maheshwar, 117
Amarendra, K., 507
Anitha Mary, Gade, 583
Ankali, Sanjay B., 623
Aparna, Giri, 583
Arunkumar, C., 407
Asha, T., 155
Asrar Ahmed, M. D., 655
Atluri, Sidharth, 377

B
Babu, B. Sankara, 475
Babu, Nirmal Varghese, 415
Badam, Nithesh, 117
Bag, Rajib, 397
Bajpai, Prasannajeet, 95
Bandyopadhyay, Aritra, 397
Banerjee, Rishin, 573
Baranidharan, V., 251
Barman, Puja, 357
Barman, Utpal, 357
Basu, Shobhin, 765
Behera, Aiswarya Priyadarsini, 715
Bej, Gopinath, 573
Benedict, Shajulin, 165
Bethu, Srikanth, 475
Bey, Marlom, 467
Bhadauria, Anuj Singh, 825
Bharadwaj, Abhinav, 17
Bharati Devi, Matta, 507
Bhateja, Vikrant, 817, 825
Bhattacharyya, Nabarun, 573, 765
Bheema Rao, N., 497
Bhima, K., 517
Bindra, Prabhleen, 609

C
Chaturvedi, Sanjay, 747
Chhabra, Indu, 701
Chourawar, Sanchita, 95
ChristyJeba Malar, A., 805

D
Dahal, Mridul Jyoti, 357
Dakulagi, Veerendra, 431
Dalal, Naveen, 701
Damodaram, Avula, 723
Darsini, R., 681
Dasari, Apurva, 441
Das, Atanu, 397
Dash, Ranjita Kumari, 681
Deb, Kaustuv, 397
Deb, Kiron, 765
Deepika, M., 345
Devi, Thongam Jayaluxmi, 185
Dey, Tamal, 573
Dhaliwal, Pankaj Preet Singh, 431
Dhana Sree Devi, K., 225
Dharan, Anandu M., 1
Donthi, Vineeth Reddy, 731

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

S. C. Satapathy et al. (eds.), *Smart Computing Techniques and Applications*,
Smart Innovation, Systems and Technologies 225,
https://doi.org/10.1007/978-981-16-0878-0

E
Eedi, Hemalatha, 773
Elaganti, Rama Lakshmi, 117
Eshwar, I. Venkata Sai, 365

F
Faruk, M. N., 271

G
Ganvkar, Nikita, 681
Gayathri, Y., 47
Ghosh, Devdulal, 573
Gopi, K., 147
Gourisaria, Mahendra Kumar, 561, 715
Gouthami, Shiramshetty, 537
Gowri, S., 207
Gujjeti, Sridhar, 241
Gundeti, Shireesha, 59
Gupt, Krishn Kumar, 609
Gurrela, Srinivas Reddy, 693
Gurunathan, Karthik, 681

H
Harikiran, J., 633
Hariprasad Reddy, Y., 9
Hazarika, Mehnaz, 357
Hegde, G. P., 487
Hegde, Nagaratna P., 287, 297, 315, 323, 345, 423, 487, 537
Hijab, Maniza, 723
Hussain, Asif Ekbal, 357

I
Indhushree, S., 261

J
Jabez, J., 207
Jacob, Meega Annu, 415
Jagan, A., 517
Jahanavi, H. S., 155
Jajodia, Nihar, 671
Jaya Murugan, R., 251
Johri, Archita, 817
Joshi, Kamaldeep, 217
Jyothish Lal, G., 833

K
Kalpana, R., 671
Kalyan, Srijha, 407
Kanade, Vijay A., 305
Kanimozhi, R. V., 261
Kar, Bimal Prasad, 127
Karthik, J., 127
Karunakara, K., 73
Kasat, Dipali, 175
Khan, Khaleel Ur Rahman, 423, 655
Khan, Pattan Asif, 773
Khonde, Shraddha R., 137
Komal Kour, S., 731
Krishna, K. Chaitanya, 365
Kshirsagar, Meghana, 609
Kshirsagar, Vivek, 609
Kulanthaivel, G., 137
Kumar, Mahendra, 431
Kumar, M. Dheeraj, 365
Kumar, Sharath, 17
Kuncharam, Manideep, 693

L
Lakshmi, D., 693
Latha, S. Bhargavi, 475
Lavanya, R., 17
Lozada T., Edwin Fabricio, 757

M
Majumdar, Sabyasachi, 573
Mallikarjuna, K., 9
Mamatha, M. N., 671
Mamidisetti, Suresh, 475
Manjunath, H., 73
Marri, Satyanarayana Reddy, 455
Menghal, P. M., 103
Miniguano Miniguano, Livio Danilo, 757
Mishra, Manasi, 549
Mukirala, Dheeraj, 731

N
Nagaraj, S., 147
Nagaraju, Animoni, 597
Nandal, Rainu, 217
Nara, Kalyani, 441
Narayan, Gowtham, 17
Narendra, Jadhav Swamini, 525
Naveen Kumar, D., 261
Navya Sree, P., 795
Nayak, Sarat Chandra, 127
Nigam, Mansi, 825
Nithya Shree, N., 251

O
Oja, Siddhi, 103

P
Pabboju, Suresh, 241
Pal, Abhra, 573
Palakurthi, Vamshi Krishna, 573, 765
Pal, Babita, 817
Pal, Deepika, 817
Pandey, Manjusha, 561, 715
Parakh, Arun, 747
Parida, Prajna Paramita, 561
Parikh, Satyen M., 333
Parthiban, Latha, 623
Patel, Ashok R., 333
Patel, Jagruti N., 333
Patel, Meghna B., 333
Prajapati, Anjali, 1
Prakash, D., 261
Pramod Kumar, P., 25
Pranathi, V., 25
Prasad, G. Lakshmi Vara, 271
Prasad, K. Lakshmi, 271
Prasad, Thulasi, 147
Pratyusha, M., 795
Prithiksha Parameshwari, A., 805
Priya Devi, Akoijam, 195
Priyamvada, R., 833
Priyanka, K. A., 155
Proddaturi, Rashmitha, 441
Pullas Tapia, Paul Santiago, 757

R
Radhika, R., 365
Raghu, Haritha, 387
Rakesh, S., 287
Rallabandi, Revanth, 377
Ramachandram, S., 655
Rama Devi, Y., 597
Ranadheer Reddy, G., 25
Rao, Sony S., 155
Rasool Md, Abdul, 643
Rasul, Shaikh Akbar Shaikh, 525
Ratkal, Bhageshwari, 59
Rautaray, Siddharth Swarup, 561, 715
Ravi, C. N., 147
Ravinarayana, B., 73
Ravishankar, Hamsini, 407
Reddy, P. Chenna, 455
Reddy, S. Pradeep Kumar, 147
Reddy, S. R. N., 549
Rekanivetha, R., 805
Rooban, S., 365
Roopa, V., 805
Ryan, Conor, 609

S
Saheb, Shaik Imam, 643
Sai Chandana, B., 633
Sai Krishna, C. R., 671
Saini, Ravi, 217
Sai Prakash Sakuru, K. L. V., 497
Sakhare, Dipti Y., 525
Sandhya, V., 315
Sarmah, Mriganka Shekhar, 467
Sarveshwaran, R., 805
Sasithra, S., 251
Satyanarayana, B., 127
Sayeedunnisa, S. Fouzia, 423
Seetha, M., 487
Senthil Kumar, G., 35
Shalini, R., 251
Sharma, Bobby, 467
Shoba Bindu, C., 225
Shobana, G., 739
Shreekumar, T., 73
Sinchana, S., 155
Singh, Khundrakpam Johnson, 185, 195
Sivasangari, A., 207
Soman, K. P., 833
Sowmya Devi, V., 323
Sowmya, V., 833
Sri Lalitha, Y., 47
Srinivasulu, Senduru, 207
Subba Reddy, T., 633
Subhash, P., 795
Subramanian, Nalini, 739
Suguna, M., 261
Sumana, Gollapalli, 583
Sumer, Jediael Meshua, 467
Suresh Reddy, Gali, 785
Suriakala, M., 665
Swami, D. R. Annappa, 73
Swathi, V. N. V. L. S., 35

T
Tattari, Jalaja, 731
Tenecota, Wladimir L., 757
Thanga Pradeep Kumar, R., 805
Thirumalai, M., 87
Thomas, Manoj V., 387

U
Udayakumar, Aparna, 387
Ulagamuthalvi, V., 137

V
Vadali, Sai Gayatri, 785
Vadivel, M., 147
Vagh, Mayuri, 175
Vaidya, Gauri, 609
Vamsi Krishna, U., 833
Vangala, Srikar, 17
Vanitha, N., 665
Varada, M. V., 387
Varija, B., 297
Vasanth, K., 147
Vasaraju, Sai Prasanna, 441
Vathsala, A. Vani, 35
Verma, H. K., 747
Vimali, J. S., 207

Y
Yuvaraj, T., 87

Proceedings of the Fifth International Conference on Inventive Computation Technologies (ICICT-2020)
IEEE Xplore Part Number:CFP20F70-ART; ISBN:978-1-7281-4685-0

An Enhanced Cluster Head Technique for Energy Consumption in Data Fusion

E.Brumancia
Research Scholar
School of Computing
Sathyabama Institute of Science and Technology
Chennai, India
easpinbrumancia@gmail.com

S.Justin Samuel
Professor
Computer Science Department
PSN College of Engineering
Thirunelveli, India
drsjustin@gmail.com

Abstract—**Wireless Sensor Network (WSN) could be presently exploitation innovation for the data collection from the stressful surroundings. WSN is consists of an outsized variety of sensing nodes, they are known as sensors. This kind of service is employed to gather the data through base stations (BS). However, it works on restricted batteries consumes a lot of loads and power, limited energy is the bottleneck. The data collected from different nodes may give redundant information. To reduce the redundant information data fusion is needed, which will increase the accuracy and robustness of the information. LEACH protocol is one of the typical data fusion algorithms, which will reduce the consumption of energy and extends the network's life cycle. The findings of the study were analyzed and compared with the well-known LEACH and GSTEB protocols. The analysis shows a significant energy-potency increase within the WSN. The results show the overall network energy consumption by scrutiny with existing algorithms.**

Keywords—WSN-Wireless Sensor Networks; LEACH-Low Energy Adaptive Clustering Hierarchy; GSTEB, Data Fusion

I. INTRODUCTION

[1]Wireless is consists of thousands restricted batteries operated nodes. Nodes are unit created of radio transmission and receiving, sensing, storage, processing, mobilize and world positioning system (GPS). These nodes area unit localized in a massive geographic region. These nodes area units are localized in a massive geographic region. These sensors area unit gift closes them then they transform the information into electrical Signals then these signals area unit processed to urge Characteristics remotely gift space wherever sensors unit. The Sensor interconnections have thought-about as a promising tool for analyzing the physical world, utilize self-organized nodes will sense, method and communicate area unit battery high-powered. In lifetime sensor interconnections node, energy must be a main yet vital resource, wherever exhibits have a restricted characteristic. So, it is often necessary to optimize specification therefore resources consumed is decreased. Their boundaries kind their style and protocols each stimulating completely different. WSNs are unit shaped by an outsized variety of sensing element.

Depending on the sensors with limited battery power, balanced energy exhausting should not be superimposed due to the reason for managing energy efficiently. The current system has been enforced by the not high Energy adjective bundle Hierarchy (LEACH) theme permits role of collection head evolved among sensing element and makes an attempt to distribute energy usage though all buds. It utilizes irregular rotation native base equally distribute energy load among senses within the track though out set-up part, all nodes are unit at instant buds through human action with short messages one lump is chosen as a bunch head per bunch head choice rule. A routing protocol (GSTEB) for the General Self structured Tree-Based Energy Balance is proposed, which have a tendency to take into account a scenario within which interconnections collects info sporadically from parcel wherever every node regularly senses atmosphere and sends data back to central Station. Commonly part unit two definitions for lifetime interconnection are time from the beginning of interconnections operation to the death of the primary node within interconnections and also the time from the start of interconnections operation death of the last node within interconnections.

The main aim of GSTEB is to do attain extended interconnections a time period various applications. Every spherical, Base station (BS) allows a root hub and communicates its ID and its coordinates to all or any device nodes. Then interconnections measure trail either by transmitting trail data or BS to system nodes or by dynamically and independently designing the same tree structure by each node. For each case, GSTEB will amendment foundation reconstruct tree with a slight delay with reduced energy exhaustion. Therefore, according to the protocols, a balanced charge is achieved. GSTEB's function is part of the initial portion, part of Tree-Construction, part of self-organized aggregation and transmission of information, and part of the exchange of knowledge.

This article is structured as following, Section II contains the Literatures, section III reviews about routing protocol, section IV contains working principles, section V contains results and discussion and section VI contains the conclusion.

II. RELATED WORK

In[2] Hongjun Wang, Huiqing Yangtze, Hui Zhao, and YoujunYue projected a formula to balance energy exhaustion of detector nodes in WSN, in order network life additionally gets extended. [3]Zhang Shiyue, Jiande Wu, Xiaodong Wang projected LEACH Cluster

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

routing formula for wireless detector Network information Fusion thereby reducing the energy exhaustion and quantity of knowledge communication. In[4] Wei-Qi bird genus; Bin- Jie Hu projected fusion formula with the gauss distribution characteristic chart, therefore, the gauss membership operate will improve the information preciseness and measuring stability. In[5] Diksha.J. Ramteke, Amol Boke projected a way supported ICH- LEACH wherever there'll be hierarchy which is able to encompass Master and slave bunch head thus on enhance the life of Master bunch head and lifelong of a specific network. In[6] the developing enthusiasm in wireless sensor networks and emerging trend in recent field motivated for measuring the characteristics, operations and communication protocols for different areas. In[7] proposed the Low Energy Adaptive Clustering Hierarchy to distribute the energy evenly between the sensors in the network. Data fusion is incorporated with routing protocols to reduce the information content that could be transmitted to the base station. Lifetime also gets extended. In paper [8] redundant data can be eliminated, which will improve the performance. So that energy consumed and network lifetime also increased. All these done by Improved LEACH algorithm. In[9] Measurement stability and data precision are improved by applying gauss membership function of fuzzy theory and improved leach protocol.

III. HIERARCHICAL ROUTING PROTOCOL

A. Leach

a) LEACH is a simple routing protocol used to reduce energy exhaustion data improves lifetime interconnections nodes. The buds in the LEACH plus cluster buds and bunch head to polling are minimum. The clustering optimization problem in single-hop networks: are used to predict lifetime reliability wireless sensor interconnections. Small energy exhaustion data report of the distance closer to base clustering on large scale. The distance from the base station far smaller cluster. Every cluster divides into mini clusters and they are organized for each mini-cluster define the other mini bunch head. It communicated directly with the bunch head and collects the information about the other nodes to connect and concentrates all nodes for data transmission used resolves some nodes coverage problems.

[10] Each cluster head is having a responsibility to create and manipulate a TDMA schedule during the communication process. To transmit the data packets program table is used by its cluster element nodes. If an emergency happens in the monitoring environment the sensing elements are activated to send the information by themselves to their own cluster head rather than the remote base station. The data from different nodes are collected by the cluster head and remove the redundant data, and then transmits the remaining information to BS. The transmission distance between the nodes will be reduced by the member node with the help of cluster head. Because of this energy consumed.

In [11] to reduce energy consumption in WSN, data fusion is the best solution. Due to sensor failure interruption may occur, by using data fusion techniques the redundant data can be reduced. By doing this entire network energy will be consumed, so that lifetime of that entire network will be increased.

GSTEB and LEACH is the best suitable algorithm to perform energy consumption.

The communication process has two different phases such as Setup phase:

[12] The node that is used to relay the message to its surrounding sensing node, the amount of which is greater than the threshold number, must pick itself as a bunch head. A node from different bunch heads will receive more than one broadcast message, and will also determine its distance by the frequency of the broadcast message received. So, buds whose number is lower than the threshold will only send a request message to the cluster containing their Broadcast ID. Which has the strongest power of the signal for energy-saving it records the nodes ID. The CH creates a TDMA time slots and broadcasts it to all the group members and each node gets its idle data transmission slots, and the steady state phase begins.

a) Steady-state phase

Cluster nodes will form them and send to BS. The formation on the bunch head will reduce unnecessary bandwidth costs and network traffic as well as a constructive indicator of energy exhaustion where the data transmission gap is shorter compared to the separate transmission to BS, which is used to save energy for the group of nodes.

A. GSTEB

GSTEB protocol is mainly proposed for wireless sensor networks the operation will be divided into phase. They are the initial phase, [13] Tree constructing phase, Self-Organized data collecting and transmitting phase and information exchange phase.

a) Initial phase

[14]In the first stage of GSTEB operation data packet to is broadcasted to all the node or few nodes by BS to share the data created time. Every node sends its data to a particular radius. Nearest node's information will be available with the packet. So that packet will be received by its nearest node. It also receives and records the data info in memory. It operates in a round, each round generates the data for BS and receives the information of nodes, once all the sensors information is received by the BS then round will be finished.

b) Tree constructing phase

The root node will be selected by the BS and broadcast the root id to coordinate with other nodes. Each and every round highest residual energy node will be selected as the root node. So the data will be collected by the particular root node and transmits the fused data to the BS. The nodes will choose the parent and its neighbors using vitality. Repeating this process until the tree is constructed. After constructing the tree, sensors will collect the information and transmits the fused data. During the transmission state, each child node sends their root id to its parent by choosing them considering their distance.

c) Information exchange phase

In every round, each and every node should generate and transmits a data packet, before its energy comes down. It can persuade the topology. So the node which is about to die should share their data with another node. This process is

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

performed based on the time slots, in each time slot random delay will be computed by the node whose energy will certainly be exhausted. This will help to make another broadcasting node in a new time slot. This particular node will broadcast the data packet to the complete network. Mean time remaining nodes will monitor the channel. Root id will be verified based on the data packet received from the node. In order to transfer the information securely the cluster head will be selected based on the degree of energy.

IV. WORKING PRINCIPLE

GSTEB is hierarchical routing algorithm which is most popular for sensor networks. GSTEB sensor nodes are combined in order to form clusters on the basis of received signal strength. The cluster head will act as routers to the sink node. [14]Transmission can be only done by bunch heads instead of all other nodes which will save energy. Existing routing protocol considers balanced node energy exhaustion ignoring extensive data processing between the nodes of the sensors. LEACH protocols need base stations with uneven distribution of bunch heads. Since the sensor consumes high energy. This should be seen as an energy-efficient management issue. The observation of energy level over sensor network provides as issues in the transmission.

The methodology performs the detector network and selects bunch heads space to the bottom station. The bunch heads are chosen by comparison their crucial values, rather than the crucial values, rather than the standard methodology of choosing bunch heads, during this bunch formula within GSTEB will perform knowledge fusion. The method will acquire higher convergence and price lower message overhead. Clustering in sensing aspect network is an economical good topology that encourages planning energy-efficient protocols to develop sensing element buds with low price.

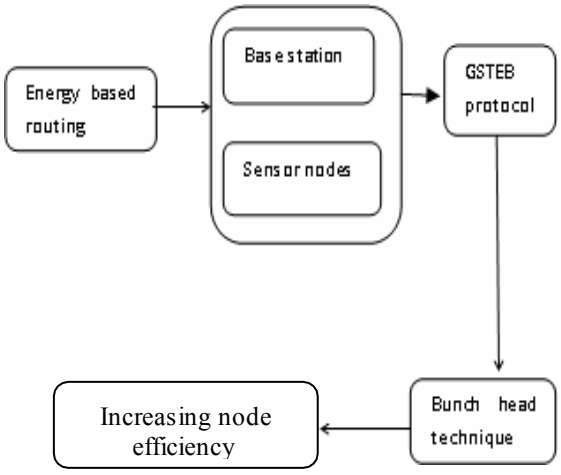


Fig. 1. GSTEB Architecture

A. Low Energy Adaptive Clustering

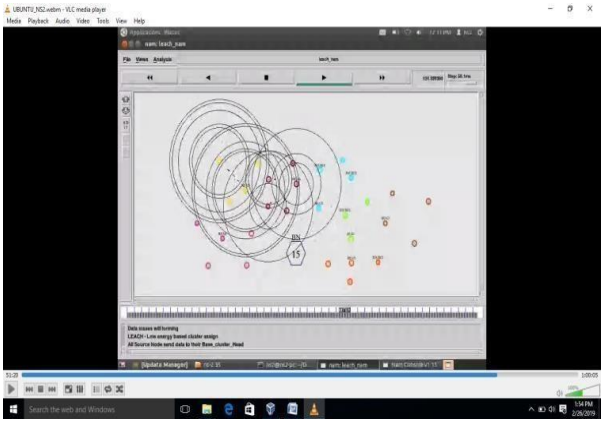


Fig. 2. Source node data to their basehead

LEACH is a simple efficiency routing protocol in WSNs. It is integrated with clustering is to lower the energy exhaustion for creating the nodes. It maintains the clusters to improve the lifetime. The performance of LEACH is slightly lower than the proposed method GSTEB.

LEACH will divide the Wireless Sensor network into a number of clusters. Each and every cluster will have a cluster head. The cluster head is responsible to collect all the data from the other sensing nodes which are shown in fig 2 and fig 3.

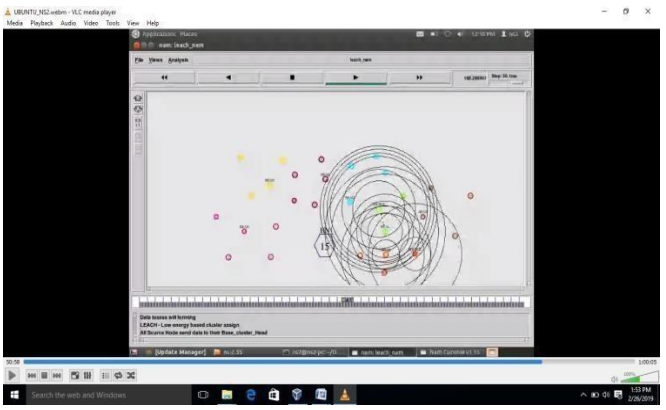


Fig. 3. Implementing LEACH algorithm

B. Tree Based Energy Balance Routing Protocol

G GSTEB is a popular energy balancing routing protocol. For each bunch head process, the base station purposes a node for each bunch head and own its neighboring ID information which node has highest weight. It is mainly used to balance the energy exhaustion and load.

STEPS OF LEACH ALGORITHM

Step 1: Begin and construct WSS and WSN
Step 2: Send node locations and residual energy information from nodes to base station
Step 3: if Node Energy> Threshold Energy then Wait for cluster
Step 4A: Send join request to CH and wait for TDMA schedule for head announcement

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

Step 4 B: Node is selected as bunch head.
Step 5: Search for clusters in interconnections
calculates the modularity (Q) of the clusters
Step 6: if Q>Qmax then
Calculate the centroid of interconnections and its distance from the nodes closest to it
Step 7: Assign highest weights to the nodes close to the centroid having residual energy than others.
Step 8: Compute the estimated (normalized data accuracy) for the sensor nodes in interconnections. Plug-in the weights of the nodes in terms of Energy and Precision constraints.
Step 9: In the weights matrix obtain the best set of cluster.

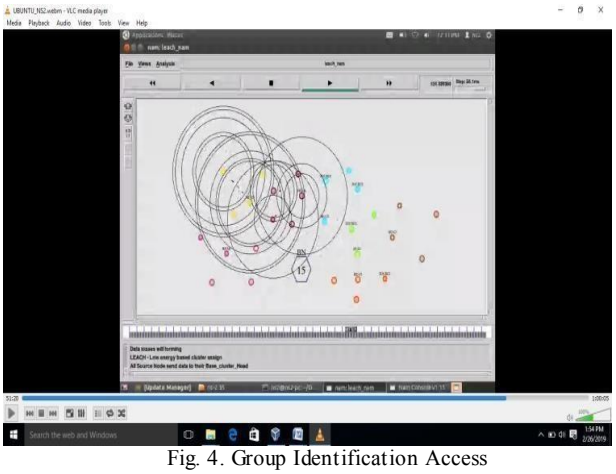


Fig. 4. Group Identification Access

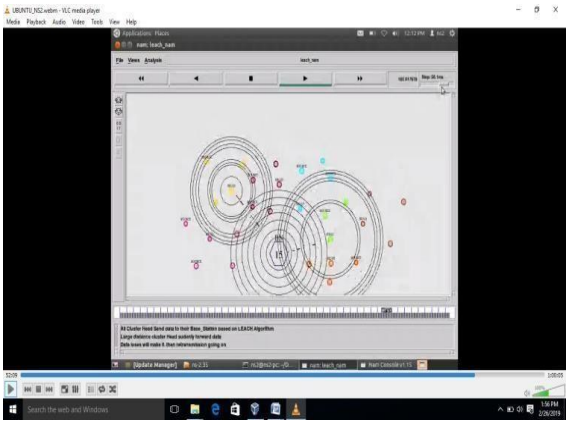


Fig. 5. Bunch head send data to bunch head

STEPS OF GSTEB ALGORITHM

Step 1: Create node
i=nodeid;
Set the routing channel;
Set inienergy, resenergy;
Set rad_range as default;
Step 2: If((node in rad_range) && (next hop!=NULL) then
Capture data load (node all);
Step3:Create a node and configure
(rreq,rrep,tsend,trev,tdrop, inienergy, resenergy)
Step 4: For i=0 to n
con_energy [i]=inienergy[i]-finalenergy [i]
total_energy[i]=<con_energy [i])
Step 5: if(max_energy[i] <con_energy[i]) then
max_energy = con_energy[i]
nodeid=i;
step 6: If (distance<300 &&max_energy [nodeid] >energy_neighbor [nodeid]) then
step 7: Max energy node is the cluster head;

V. RESULT AND DISCUSSION

In LEACH, based on the distance the sensor energy consumption was reduced. The nodes will concentrate all the nodes. It only concentrates on specific nodes that transmit the data and information to the other node. It will load the data nodes more and more as shown on figure 4 and figure 5.

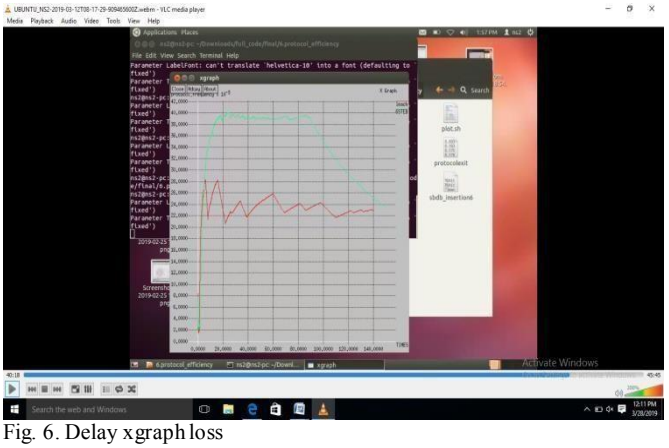


Fig. 6. Delay xgraph loss

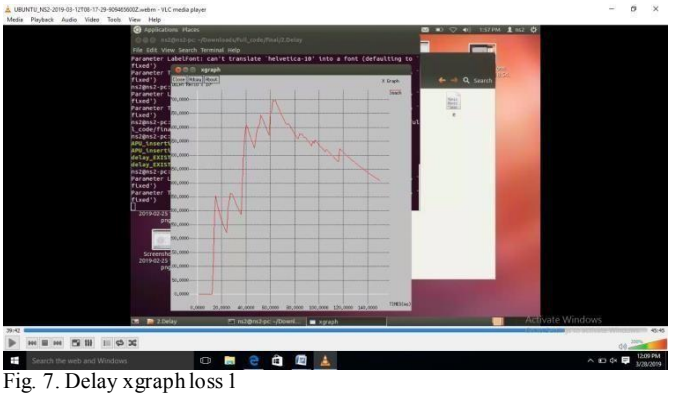


Fig. 7. Delay xgraph loss 1

In GSTEB, the nodes will concentrate all the nodes and do the data processing even though bandwidth is high and minimize energy consumption efficiently compared to LEACH on shown in figure 6 and figure 9. It reduces load and have used E-GSTEB. The type of GSTEB is for minimizing energy consumption

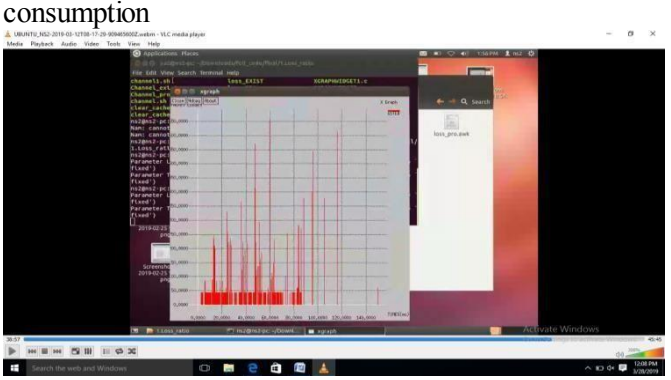


Fig. 8. Parameter xgraph loss

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

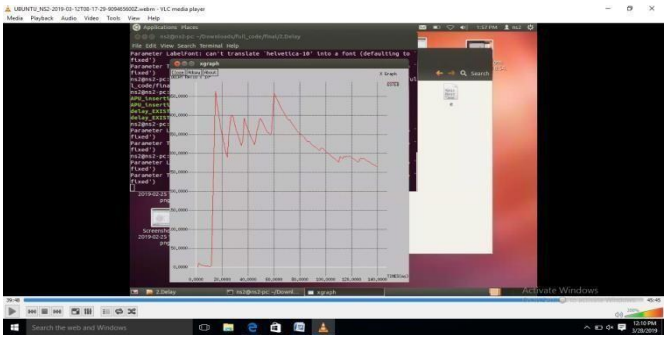


Fig. 9. Parameter xgraph

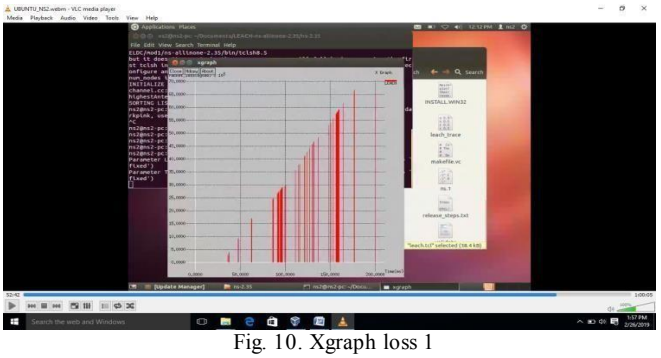


Fig. 10. Xgraph loss 1

The Base station allocates nodes coordinates in each round. Interconnections transmit the data information being dynamically and individually

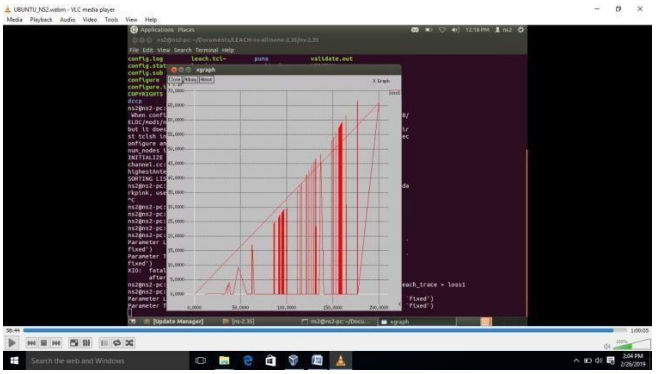


Fig. 11. Xgraph loss 2

The GSTEB chooses the nodes with residual energy to transmission the data information. The allocation of nodes will performed by the Base Station based area unit distance with the radius clusters are shaped. All the nodes will be covered with the other nodes.

VI. CONCLUSION

In this paper, an enhanced cluster head technique for energy consumption in data fusion is proposed. Data from different sensors are grouped and performed cluster heading. The simulation results show the minimized energy with enhanced (E-GSTEB) technique. Because of this energy reduction, network lifetime will get improve. This clustering technique gives more accurate data and data transmission is more efficient. This is done by using a particular data fusion technique.

ACKNOWLEDGMENT

We're thankful to our management Sathyabama Institute of Science and Technology for providing us with all the resources and facilities needed to prepare this manuscript.

REFERENCES

[1] Junguo Zhang, Wenbin Li, Xueliang Zhao, Xiaodong Bai, Chen Chen, "Simulation and Research on Data Fusion Algorithm of the Wireless SensorNetwork Based on NS2", 2009 World Congress on Computer Science and Information Engineering

[2] Hongjun Wang *1 and Huiqing Chang*1, *2Tianjin Agricultural University," Research on LEACH Algorithm Based on Double ClusterHead Cluster Clustering and Data Fusion", Proceedings of 2017 IEEE International Conference on Mechatronics and Automation, August 6-9, Takamatsu, Japan

[3] Zhang Shiyue, Jiande Wu, Xiaodong Wang "A Data Fusion Algorithm Based on Improved LEACH Clustering Routing for Wireless Sensor Network", icaise.2013.7

[4] Wei-Qi Chen,Bin-Jie Hu, "A Data Fusion Algorithm in Leach Protocol Using Gauss Membership Function", 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM) 23-25 Sept. 2010

[5] Diksha J. Ramtekel , Amol Boke2, "Modification of LEACH Protocol in Wireless Sensor Network: A Review," International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 05 | May-2018

[6] K. Akkaya, and M. Younis, "A survey on Routing Protocols for Wireless Sensor Networks," Elsevier Ad Hoc network Journal, vol.3, pp.325-349, 2005.

[7] W. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy efficient communication protocol for wireless microsensor networks," IEEETransactions on Wireless Communications vol. 1 (4), pp. 660–670, 2002.

[8] Zhang Shiyue, Jiande Wu, Xiaodong Wang, "A Data Fusion Algorithm Based on Improved LEACH Clustering Routing for Wireless Sensor Network", August 2013

[9] Wei-Qi Chen ; Bin-Jie Hu "A Data Fusion Algorithm in Leach Protocol Using Gauss Membership Function", 2010 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM), 14 October 2010

[10] Fuzhe Zhao, You Xu, and Ru Li," Improved LEACH Routing Communication Protocol for a Wireless Sensor Network", International Journal of Distributed Sensor Networks Volume 2012, Article ID 649609, 6 pages

[11] E. Brumancia, S. Justin Samuel, M. Gomathi and Y. Mistica Dhas, "AN EFFECTIVE STUDY ON DATA FUSION MODELS IN WIRELESS SENSOR NETWORKS", ARPN Journal of Engineering and Applied Sciences, VOL. 13, NO. 2, JANUARY 2018.

[12] SudhanshuTyagia, NeerajKumarb, " A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks", Journal of Network and Computer Applications Volume 36, Issue 2, March 2013, Pages623-645

[13] M.Sengaliappan1 , Dr.A.Marimuthu2, "Improved General Self-Organized Tree-Based Routing Protocol For Wireless Sensor Network", Journal Of Theoretical And Applied Information Technology 10th October 2014. Vol. 68 No.1

[14] Manjinder Kaur Dr. Shashi B. Rana*, Performance Comparison of Leach and GSTEB in Wireless Sensor Network, International Journal of Engineering Research & Technology (IJERT), Vol. 4 Issue 04, April-2015

[15] Yan Qiang,1 Bo Pei,1 Wei Wei,2 and Yue Li, "An Efficient Cluster Head Selection Approach for Collaborative Data Processing in Wireless Sensor Networks", Hindawi Publishing Corporation International Journal of Distributed Sensor Networks Volume 2015, Article ID 794518, 9 pages http://dx.doi.org/10.1155/2015/794518

[16] E. Brumancia, S. Justin Samuel, M. Gomathi and Y. Mistica Dhas, "AN EFFECTIVE STUDY ON DATA FUSION MODELS IN WIRELESS SENSOR NETWORKS", ARPN Journal of Engineering and Applied Sciences, VOL. 13, NO. 2, JANUARY 2018.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

[17] B.Warneke, K.S.J. Pister, "MEMS for Distributed Wireless Sensor Networks," in Proc. of 9th International Conf. on Electronics, Circuits

[18] K. Sohrabi, et al., "Protocols for Self-organization of A Wireless Sensor Network," IEEE Personal Communications, vol. 7, No. 5, pp. 16-27, October, 2000.

[19] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey On Sensor Networks", IEEE Communications Magazine, vol.40, pp.102-114, 2002.

[20] Sudhanshu Tyagi and Neeraj Kumar, "A systematic review on clustering and routing techniques based upon LEACH protocol for wireless sensor networks," Journal of Network and ComputerApplications vol. 36, pp.623–645, 2013.

[21] Zhao Han, Jie Wu, Jie Zhang, Liefeng Liu, and Kaiyun Tian, "AGeneral Self-Organized Tree-Based Energy Balance Routing Protocol for Wireless Sensor Network," IEEE Transactions On Nuclear Science, vol. 61, No. 2, April 2014.

[22] Dheeraj and Ritu Mishra, "Review Paper on Hierarchal Energy-Efficient Protocols in Wireless Sensor Networks," International Journal of Advanced Research in Computer Science and Software Engineering, vol.4, Issue 6, 2014.

[23] S. Upadhyayula, S.K.S. Gupta, "Spanning tree based algorithms for low latency and energy efficient data aggregation enhanced converge cast (DAC) in wireless sensor networks, Ad Hoc Networks, vol.5, pp.626–648, 2007.

[24] Andrei Gagarin, Sajid Hussain, Laurence T. Yang, "Distributed hierarchical search for balanced energy exhaustion routing spanning trees in wireless sensor networks," J. Parallel Distribution Computer, vol.70, pp.975-982, 2010.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on February 01,2023 at 06:48:09 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

An Enhanced Adaptive Random Sequence (EARS) Based Test Case Prioritization Using K-Mediods Based Fuzzy Clustering

N.Gokilavani[1]
Research Scholar, Dept. Of computer science and engineering, Sathyabama institute of science and technology, Chennai, India
gokilavani_testing@yahoo.com

Dr.B.Bharathi[2]
Professor, Dept. Of computer science and engineering, Sathyabama institute of science and technology, Chennai, India
bharathi.cse@sathyabama.ac.in

***Abstract*—The efforts of prioritization method is to maximize the detection of fault rate by organizing the significant test cases which is operated in a sequence of regression tests. Generally, it is implemented to sort down the test cases based on the priorities former than those with minimum priority imparting to an estimated criteria. The faults which gives maximum impacts should be detected at earlier stages in testing practices. The adaptive random testing is implemented to execute arbitrary testing through input triggering clustering errors. It improves the detection ratio of regression testing in software based on object-oriented. In this proposal, an adaptive techniques of test case prioritization relied on fuzzy clustering is implemented. The adjacent matrices is generated and cluster head is chosen within the test cases. It is made by identity precise pairing. Then Enhanced Adaptive random sequence depending on prioritization of test cases detects the flaws which operates to categorize neighboring test cases as varied as possible. Hence the outcomes proved increased efficacy in earlier fault detection rate.**

Keywords—Adjacency Matrix; Fuzzy k-medoid; Adaptive Random Sequence (ARS); Average Percentage of Fault Detected.

I. INTRODUCTION

The growth in software development requires a sustainable regression to manage the high quality in the current competitive circumstances for testing operation. Nowadays, the prioritization in the test cases is the modern progressive method to compensate for the enhancement requirement of software testing efficacy. [1]This method is better for earlier fault identification, and so the debugging can be done at more initial stages. The present testing method decides implementation patterns of all test sets depending on the attention of either loop or coding on behalf of project and product capacity to eliminate the faults. It doesn't aim at the execution of data for prioritization of test cases.

A. Predicted issues

Even though there are various methods suggested for test case prioritization, but still, there are few difficulties when it is applied in a real-time environment. The imperfection in the software is occurred due to the sparse coding, which should be identified and rectified. The significance of the identical test cases should be recognized that denotes the functional dependencies. The active test case prioritization should enclose multiple criteria such as fitness values, fault data, code coverage, and historical execution of data.

B. Scope of the research

In the software development lifecycle, software testing is a mandatory process which should be performed effectively. A focused approach should be insisted to determine the precipitate optimum glitches by considering the test cases in a detected way. This method of generating test case is referred to test case prioritization. In this study, the adaptive clustering approach is developed that executes based on the statement level fuzzy optimization methods. The main objective of this research is to identify the bugs as the principal element as conceivable. The test case prioritization recognizes the similarities of evolution in regression testing by applying a fuzzy-based clustering method. Generally, software testing is a time-consuming task and expensive to be used. So, it is necessary to design the solution in optimal cost for declining the testing methods. It should be calculated how the priority wise test cases varies from the non-prioritize test methods. Because the test case of non-prioritize documents will run by test procedure and the further step would be the testing phase. In sequential test cases, the drawbacks of testing circumstances is fed as input for any of the prioritization algorithm. Conferring to the mentioned priorities, the testing case will be accomplished.

[2]The adaptive approach gives the required data concerning the current execution of a application at execution time. It computes the efficiency of the test cases, which collects the data flaws and organize the priority of every situation in subsiding order. [3]The adaptive approach selects the test case and performs once the output of the process is achieved when the priority of the test cases is insisted. One by another produces the real arrangement of the test case. After execution, the fresh order is created based on priority. The original test case order focused on the priority test case order, which is not matched. Since it has less faulty test cases which are marked as the first number of the sequences which is not possible by original test cases since it's prone to more attacks. There are several metrics which are evaluated with the aid of prioritization techniques (test case). [3]It is implemented for programming the order of test cases. The adaptive algorithm is processed to give the prioritization. In this research, the Enhanced Adaptive Random Sequence

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

(EARS) is implemented to organize the test cases. The programming of test cases is very noteworthy for completing the increased amount for test case. The main purpose of this research is to detect defective modules. In this proposed system, finding fault in test cases will be computed at the selection of test cases. Initially, chose a minimum test for implementation. When the first test case is completed, the performance will be stored as its output. The selection process is persistent until test cases are completed. Few statements are not selected because they will not get executed. The adaptive method is best strategies for test case prioritization. The novelty in this research is the enhanced adaptive algorithm is as follows,

- The adjacent matrix is produced and listed relied on the flaws of test cases. So the method selects the candidate who owns the increased faults as cluster head.
- The fuzzy clustering technique, along with integrated k-medoids, is implemented for creation of clusters with identical case attributes. It is because of the cases in the same group and is diverged in commencing of preceding auxiliary clusters.
- The enhanced adaptive random sequence is abbreviated as EARS used to enhance the failure identification by continually spreading the test input overall finishing the input province. The proposed algorithm adds up in the choosing of the ordered test case as divergent along the input region as possible.

This research paper is organized into several sections as below. Part II explains the representation of a related survey for clustering methods and test case prioritization methods. Section III elaborates the description of the proposed method in selection of optimal path and cluster head and a way for Test case prioritization. Section IV explains the results, discussion and comparison with other methods. Section V concludes the research with analysis on the future extension for the proposed method

II. REVIEW ON RELATED WORKSFGDGD

The different techniques implemented to rectify the shortcomings of test case prioritization is discussed in this section. The positive and negative aspect of various fuzzy clustering methods based on the test cases is deliberated to overcome its limitations.

A.Test case prioritization

This part reviews the existing methods developed in the past decade on test case prioritization methods. This method is very imperious in regression testing. Generally, it is used to create the order of the conforming test cases. [4]The existing test cases with high priorities are completed prior than the case with lesser priority. Comparing the optimal prioritization methods with nine sections such as no prioritization, optimal, random prioritization, branch-addtl, branch-total, FEP-addtl, FEP-total, stmt-addtl and stmt-total. The enhancement in every process concluded that improvised rate of fault detection. But limitations of research executes only with the base application rather than the modified version of an application. [5]As proposed by, the software license works on with five standard prioritization to utilize the chances of main problem recognition ratio of SPL test cases. The similarities and variations are used for computation purposes. The maximum value of the product variant is accounted as every product's minimum variation at a new system that needs an end to end retesting. [4]Revised all test case prioritization methods and factors based on client requirements, coverage, history and cost. [6]As discussed by, when there are any alterations in the client requirements, the impacts of fault are defined. Test cases are well organized according to the ethics of fault proneness and customer priorities. This work covers three optimal techniques with NSGA-II to extend both diversity and coverage for minimizing the test accomplishment duration. It is compared with coverage and justified the efficacy of the merged techniques. The disadvantages of this method eliminate the cost of unidentified faults. [7]Proposed a section of the program to execute test case prioritization as more effective. It is cost-effective, and the bugs are identified at an earlier stage. But it is not effective as it consumes more time for execution. Discovered the new method where fault detection is based on direct dependency and indirect dependency. But the size of the program is not portable, which influences the execution of test case prioritization. [8]Implemented three similarity features such as Hamming distance, edit distance and basic counting to increase the efficacy of the testing process. [9]Suggested a distinct approach consider the unselected test cases based on the recent nomination of the software's original version. The factors such as value, fp, coverage programming, and litigation and coverage granularities are improved in the prioritization methods. It gives that variable have required origin in task and test cases which should be tagged for every variable to give different clarifications and accessible attributes from variables essential for research. [10]Explained about the common causes encountered in the problem retrieval, that can localize the chaos of regression test prioritization such as the variance of documents organized between two test versions. [11]To improve the capability of earlier fault detection, proposed DIP to use the management of data and control dependence variables by sequences of collaboration but the length of the code gives complexity which should be resolved further

B. Cluster based prioritization of test case

This review deliberates the clustering-based test case prioritization methods. [12]Explained the similarity of cosine to cluster the similar values for cases (test) . The priority values are calculated by feeding the clusters as input. It is made simultaneously to limit the computation time and maximize error recognition. Proposed three different variables such as information of fault history, complexity measure and coverage of code is implemented to the clusters to build the prioritization model. [13]Deliberated to implement priorities to work on descending order with six factors of the test cases which primarily doesn't focus on clustering, but it operates on with cluster data.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

C. Test case optimization techniques based on soft computing

[14]Explained a fuzzy approach to allocate the priority by various attributes that include coverage based parametric values, variety of event and collaboration. The priority of test cases are organized depending on fitness functions, and the pattern of the test suite is accomplished. The fuzzy method gives improved result other than existing methods. [15]Implemented a fuzzy rationality to control the prioritization of cases. It computed the defilement of a specified application assertion. The prior data on violation history is accounted for assessment when it is applied for the event formation. Suggested a prioritization method which is based on completed similar proceedings. The events are organized to the significance of their related structures, but it is not as efficient as the following bug location affects the performance of the further process. [16]It has implemented a fuzzy control approach for identifying the priority of cases through execution trees which are utilized in common languages to define the priorities. This method prevents the effortless results in the test by the symbolic execution of trees to manage where the test gives the most advantages to execution coverage.

The reviews, as mentioned earlier, stressed the need for the effective algorithm as the existing tools don't give remarkable efficacy in optimization and prioritization. Because some approaches are code-based or component-based, which is focused test case prioritization, it is cost-effective. So the combined methods of approximate analogous cases of clustering approaches and efficient prioritization algorithm for identification of faults as early as possible. The comprehensive methods lead in faster execution of the program.

III. PROPOSED SYSTEM

The detailed view of the proposed methodology by FCAR technique is integrated with k-medoids based on fuzzy clustering along with EARS relied on test case prioritization methods. It offers a great solution for test case prioritization methods. The overview of proposed cluster creation and discourse the merits and prioritization. The building of succeeding process to attain the test case prioritization to define the faults as early as possible is as structured below,

- Preprocessing
- Finding the literal pair
- Executing an adjacency matrix
- Selection of cluster head
- Computation of similarities
- Formation of clusters
- Enhanced adaptive random sequence prioritization

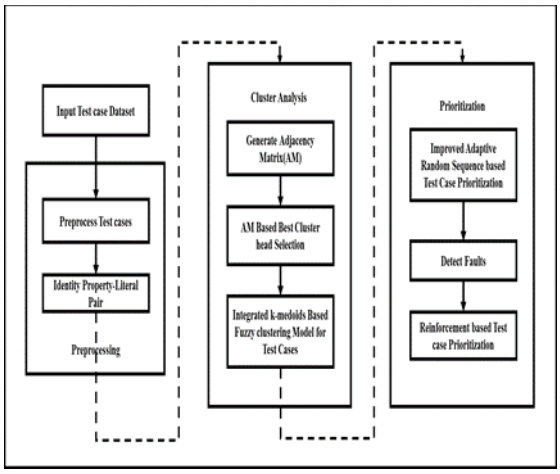


Fig 1. Overall flow of the proposed system

The fig.1 illustrates the workflow of the proposed framework of test case prioritization approach, which composed of chronological progressions. Predominantly, the test case dataset is fed as input to preprocessing block where the code and event dependences are explained in detail. Then a similar pair of test cases are developed, which is helpful to define the adjacency matrix. Then the adjacency matrix is implemented to considering every neighboring test case, where the cluster head is chosen. The assimilated cluster set may not be accounted as accurate as is a club of similar test cases. So, integrated k-medoids based fuzzy clustering techniques is executed for the creation of clusters with identical attributes of test cases. Then the enhanced adaptive random sequence prioritization method is applied to identify the faults in test cases as ordered.

A. Preprocessing and Pairing

The preprocessing of every test cases in the testing dataset is made to remove the duplicates and noises. Then the corresponding output is stored and aggregated in the testing database. The testing database load functions perform the tester delivers the testing criteria and store it in the table. The selection of test cases differs for each new product. The identity properties in the generation of test data include trail, loop, or description which is coupled for further use in the testing process. The pairing methods help in parallel processing which supports to enhance the presentation of the searches. Looping events are identified at the initial stage for the identification of the test data to be produced. The evaluation is mandatory for the recognition of diversity in the research programs. It could be reasonably multi-layered and may probably include extreme perceptions on the view of the developer and initiator of test data to define earlier possibilities in various programming scenarios. According to pairing, the test case is accounted to be useful when a pair of definition and variable usage is maximum.

B. Execution of adjacency matrix and selection of cluster head

The equivalent paring values are logged for each cases that indulges the requirements of test properties such as branch

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

coverage or statement. Generally, a test case is executed and checked on its pairing values. It is chosen and located in the adjacent matrix for every iteration, and subsequent loop or report is met by a test data. The execution of circuits and codes as predicted, if the primary value is zero, is increased by one in the generated matrix. It persists for every test cases when test cases meet the resultant report. The capacity of the model is based on generated paths. The matrix is built with several coverage and branches. The main focus of the adjacent pattern is to expose all the faults. The matrix halts when ensuring of total flaws are masked through it or not. If not, then chose the adjacent test cases that checks the remaining flaws and repetition on this whereas entire bugs have been encircled. When the whole bugs are surrounded, calculate an amount as the whole of errors enclosed using every test case placed in the adjacent matrix. This approach would guide a number of combining of test events known as coding structures that ensure that all software errors are handled. Then the test case that covers the highest faults is chosen as the cluster header.

C. Formation Of Cluster

It implemented the integrated k-medoids based on fuzzy clustering methods along with failure creating inputs. The clustering is the main progression which paved to favorable state of affair for ARS. [17]The adjacent dimensional matrix which is N*N is fed as input to a k-medoids algorithm that assigns every test case to one of C clusters to minimize the sum squares of the cluster.

$$\sum_{i=1}^{c} \sum_{k \in A_i} \| x_k - v_i \| 2 \qquad (1)$$

Here, the centre of the cluster is indicated as adjacent test cases to means of test cases in a single group, $V = \{V_i \in X | 1 \leq i \leq c\}$. A_i and v_i is the set of test cases and mean in the cluster i^{th}.

The test space in the matrix is discrete as each test case denotes the position of the adjacent matrix. The fuzzy K-medoids technique minimizes

$$J_m(V, X) = \sum_{i=1}^{n} \sum_{i=1}^{c} U_{ij}^{m} \, r(X_j, V_i) \qquad (2)$$

The minimization is attained with complete V in X, U_{ij}characterizes the fuzzy connection of X_j in cluster i. The relationship U_{ij} can be demarcated intended to increase the possibility of resolving the testing difficulties in various ways. The fuzzy membership design is represented as.

Where $m \in [1, \infty)$ is the fuzzifier.

$$U_{ij} = \frac{\left(\frac{1}{r(X_j, V_i)}\right)^{1/(m-1)}}{\sum_{k=1}^{c} \left(\frac{1}{r(X_j, V_i)}\right)^{1/(m-1)}} \qquad (3)$$

Algorithm 1 Fuzzy K − Medoids Algorithm

$Assign\ the\ count\ of\ cluster\ c;\ Set\ i\_ter\ = 0;$
$Randomly\ select\ the\ initial\ set\ of\ medoids:$
$V = \{v_1, v_2, \ldots . v_n\}\ from\ X_c;$
$Repeat$
$for\ i = 1\ to\ c\ do\ /*\ Estimate\ \ memberships\ */$
$for\ j = 1\ to\ n\ do$
$Compute\ u_{ij}\ by\ applying\ (3), (4), (5)\ or\ (6)$
$end\ for$
$end\ for$
$Record\ the\ present\ medoids: V^{old} = V;$
$Calculate\ the\ new\ medoids:$
$for\ i = 1\ to\ c\ do$

$$q = argmin \sum_{j=1}^{n} U_{ij}^{m} \, r(X_k, X_j) \qquad 1 \leq k \leq n$$

$v_i = X_q;$
$end\ for$
$i\_ter = iter + 1;$
$Until\ (V^{old} = V \quad or\ i\_ter = MAX\_ITER)$

D. Algorithm for prioritization

EARS is being used to create ARS to organize a group of well-arranged test cases as a variant of the input province as conceivable. The primary objective of the proposed technique is to define the marginalization sections closer to each test case performed.

The arbitrary test cases that belong under the marginalization parts is not preferred as test cases. It acquires new test cases which are away from the operated test cases. The arbitral values in the selected cases inside the cluster are initially checked. If it meets the requirements for the test cases, then shift to the next matters. Hence the testing time is limited to the considered level and reduce the computation overhead by increasing the variability of test cases.

Algorithm 2 Improved Adaptive Random Sequence(IARS)

$Set\ m\ = 0\ and\ S\ = \{\}$
$Unsystamatically\ choose\ a\ test\ case, tc, from\ the$
$input\ suites\ based\ on\ uniform$
$distribution.$
$Increment\ n\ by\ 1. If\ tc\ expose\ bugs,$
$go\ to\ final\ Step;$
$Else, record\ tc\ in\ S.$
$arbitrary\ form\ \ D$
$= \{d_1, d_2 \ldots d_k\}\ according\ to\ the\ uniform$

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

distribution

For each element in D, find its nearest neighbour in S.

In D, find which element, d_j, has the farthest distance to its nearest neighbor in S.

Let tc $= d_j$ and goto above increment step

Return m and t. EXIT.

IV. RESULTS AND DISCUSSION

The result of the existing methods and proposed techniques on prioritization is compared by evaluating the performance metrics such as fault detection rate and the average per cent of fault detected.

A. The Average Percentage Of Fault Detected (APFD)

It is used to predict the test case ordering rehearsals. If C is a test category closing in test cases, k then B will be a set of bugs p discovered by C. The initial test case index is CBi which organizes the C ' of cluster C and shows the problem i.

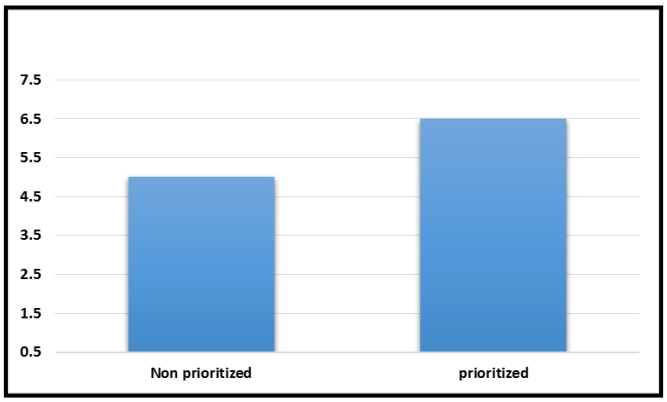


Fig 2. Prioritized VS Non Prioritized Test Cluster

The result of AFPD is associated for both prioritized as well as non-prioritized test cases. The result insists that prioritized cases deals earlier faults than one with non-prioritized cases.

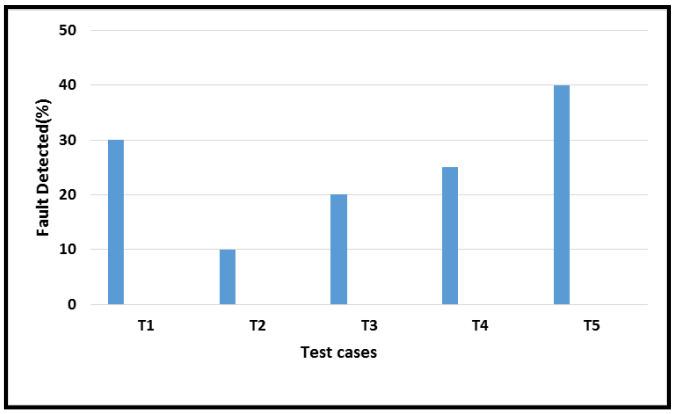


Fig 3. APFD of each test cases

The percentages for detection of faults is based on the evolution of given test order such as T4, T2, T1, T5 and T3 which gives the value of 0.65 whereas the prior arrangement shows 0.42

V. CONCLUSION AND FUTURE WORK

The proposed method is based on an improved cluster based fuzzy algorithm, which deliberates the understanding of medoids. The difficulty of this technique is O (n2) for every iteration at the time of imparting. The complexity narrates the positivity with the future fuzzy algorithm. The pairing idea offers coverage depends on the best path by the construction of the particular adjacent matrix for every iteration. The kind of test cases which demonstrates many faults is chosen as cluster head. The integrated k-medoids depend on on fuzzy clustering gives the preeminent test case when a similar error occurs. Then EARS algorithm detects the priorities to make organized and listed sequences. Hence the first result proved effective performance in software testing. In future, the proposed method is extended in high dimensional input to achieve increased performance metrics.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g.” Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] S. Elbaum, A. G. Malishevsky, and G. Rothermel, "Test case prioritization: A family of empirical studies," *IEEE transactions on software engineering,* vol. 28, pp. 159-182, 2002.

[2] J. Chen, L. Zhu, T. Y. Chen, D. Towey, F.-C. Kuo, R. Huang*, et al.*, "Test case prioritization for object-oriented software: An adaptive random sequence approach based on clustering," *Journal of Systems and Software,* vol. 135, pp. 107-125, 2018.

[3] B. Jiang, Z. Zhang, W. K. Chan, and T. Tse, "Adaptive random test case prioritization," in *2009 IEEE/ACM International Conference on Automated Software Engineering*, 2009, pp. 233-244.

[4] Z. Sultan, S. N. Bhatti, R. Abbas, and S. A. A. Shah, "Analytical review on test cases prioritization techniques: An empirical study," 2017.

[5] A. B. Sánchez, S. Segura, and A. Ruiz-Cortés, "A comparison of test case prioritization criteria for software product lines," in *2014 IEEE Seventh International Conference on Software Testing, Verification and Validation*, 2014, pp. 41-50.

[6] D. Mondal, H. Hemmati, and S. Durocher, "Exploring test suite diversification and code coverage in multi-objective test case selection," in *2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST)*, 2015, pp. 1-10.

[7] C. Fang, Z. Chen, K. Wu, and Z. Zhao, "Similarity-based test case prioritization using ordered sequences of program entities," *Software Quality Journal,* vol. 22, pp. 335-361, 2014.

[8] C. P. Rao and P. Govindarajulu, "Automatic Discovery of Dependency Structures for Test Case Prioritization," *IJCSNS,* vol. 15, p. 52, 2015.

[9] T. B. Noor and H. Hemmati, "A similarity-based approach for test case prioritization using historical failure data," in *2015 IEEE 26th International Symposium on Software Reliability Engineering (ISSRE)*, 2015, pp. 58-68.

[10] D. Hao, L. Zhang, L. Zhang, G. Rothermel, and H. Mei, "A unified test case prioritization approach," *ACM Transactions on Software Engineering and Methodology (TOSEM),* vol. 24, pp. 1-31, 2014.

[11] C. N. A. Petrus, M. Razou, M. Rajeev, and M. Karthigesan, "Model-Based Test Case Minimization and Prioritization for Improved Early Fault Detection Capability," *International Journal of Innovative Technology and Exploring Engineering (IJITEE),* vol. 2, pp. 205-210, 2013.

[12] R. Kanimozhi and J. Rbalakrishnan, "Cosine similarity based clustering for software testing using prioritization," *IOSR Journal of Computer Engineering (IOSR-JCE),* vol. 16, pp. 75-80, 2014.

[13] M. H. DR, "Clustering approach to test case prioritization using code coverage metric," *International journal of engineering and computer science,* vol. 3, 2014.

[14] N. Chaudhary and O. Sangwan, "Assessment and Comparison of Fuzzy Based Test Suite Prioritization Method for GUI Based Software," *Assessment,* vol. 7, 2016.

[15] A. M. Alakeel, "Using fuzzy logic in test case prioritization for regression testing programs with assertions," *The Scientific World Journal,* vol. 2014, 2014.

[16] E. J. Rapos and J. Dingel, "Using Fuzzy Logic and Symbolic Execution to Prioritize UML-RT Test Cases," in *2015 IEEE 8th International Conference on Software Testing, Verification and Validation (ICST)*, 2015, pp. 1-10.

[17] R. Krishnapuram, A. Joshi, and L. Yi, "A fuzzy relative of the k-medoids algorithm with application to web document and snippet clustering," in *FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No. 99CH36315)*, 1999, pp. 1281-1286.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:16:28 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)
IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2

Air Pollution Detection and Prediction Using Multi Sensor Data Fusion

E. Brumancia, S. Justin Samuel, L.Mary Gladence
easpinbrumancia@gmail.com

***Abstract*— The level of air pollution is rising day by day which affects human life and the environment. Hence, it is important to control it by keeping an aircheck at its level constantly. To reduce the pollution level we develop a multi-sensor data fusion model that detects and predicts the most severe gas. In this work we suggest an effective algorithm to cluster the multiple sensors data which used to group the data and partition the data. This paper proposes a random forest approach to predict air quality. The efficiency is measured with real multiple city-data.**

***Index Terms*— Multisensor, data fusion, air pollution, prediction, Linear Discriminant Analysis, Random Forest**

I. INTRODUCTION

Air pollution is a combination of organic and man-made pollutants in the air we breathe. It is typically classified into two divisions: external air pollution, and internal air pollution. External air pollution entails exposures to un-built ecosystems. For example, Pollution from such sources as cars and power stations, factories, etc are polluting the air. Whereas indoor air pollution is from household products, chemicals, building materials, tobacco smoke, gases, etc. The air quality control system has been used to monitor air pollution levels. Over the past 30 years, studies have discovered a wide range of health consequences that are thought to be linked with air pollution exposure. Some air pollutants are harmful. Inhalation of these pollutants can raise a chance of having problems with your health. People with lung or heart problems, Children's and late teenagers are at a greater risk of emissions from the atmosphere. It doesn't mean we have pollution only outside, the air inside buildings can also be toxic and harm your health. Air pollution is defined as the presence of chemical components known as pollutants in the air. The release of chemicals into the atmosphere leads to air pollution. These pollutants change the composition of the air, making it hazardous. The main source of air pollution is fossil-fuel burning as fossil fuels are burned a lot of chemical components are released in the air which causes air pollution. Air pollution has its implications on humans and the environment. This causes human respiratory and heart problems. When the polluted air is inhaled by humans the organs are affected. The other effects are global warming, industrial pollution, eutrophication, ozone depletion, etc. Air pollution can be reduced by using public transport, using energy-efficient devices, using solar energy and wind energy, etc. Using public transport instead of private transport reduces the emission of pollutants into the atmosphere. Using energy-efficient devices and renewable resources like solar power geothermal power can lower the consumption of electricity produced by burned fossil fuels.

Poor air quality is killing the people. The World Health Organization states Poor outdoor weather in 2016 caused an estimated 4.2 million premature deaths, In low- and middle-income nations, about 90 percent. Three billion citizens who are cooking or heating their homes by burning wood, coal, and gasoline get health problems due to indoor smoke. Higher rates of air pollution have been linked with a tumor, cardiac disease, stroke, and breathing disorders like asthma. Around 134 million Americans, according to figures by the American Lung Association. Owing to air pollution and premature death, about 40 percent of the population is at risk of disease. Major air pollutants are particulate matter, lead, sulfur dioxide, asbestos, carbon monoxide. Particulate matter smaller than 10 micrometers poses a higher risk to health, breathing deep into the lungs, then moving through the bloodstream.

The integration of multiple data sources to produce more consistent data is known as data fusion. Sensor fusion is defined as the combination of multiple sensors data to provide descriptions about an environment. Sensor fusion which is a subset of information fusion is also called Multisensor fusion. The multi-sensor data fusion can be further classified into homogeneous data fusion and heterogeneous data fusion. Homogeneous data fusion is the fusion of data of the same type of sensors and heterogeneous fusion of data integrates multiple sensor data. The different levels of data fusion processing are a fusion of the signal level, a fusion of the object level, a fusion of the function level, and fusion of the decision level.

Air pollution sensors are devices used when measuring air pollution. Those sensors track and detect pollutant concentrations in the air. These sensors majorly focus on five components, they are ozone, particulate matter, sulfur dioxide, carbon monoxide, and nitrous oxide. These are very costly sensors when released, but with technological growth, they have become affordable now.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

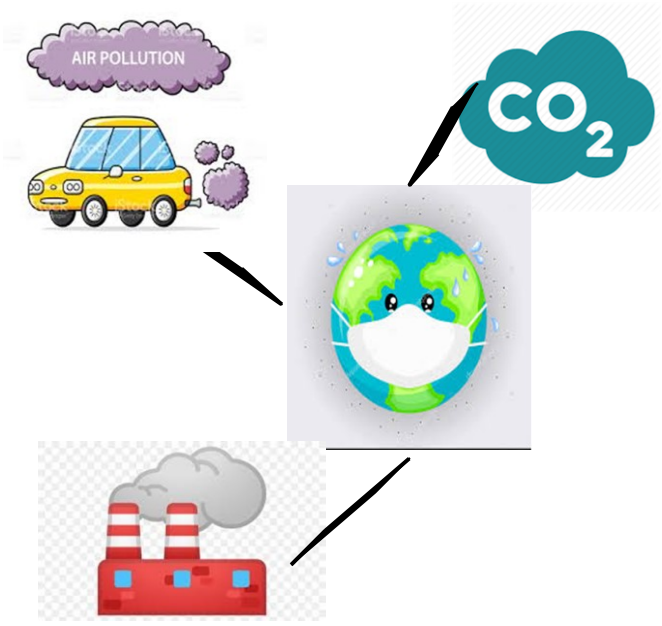


Fig. 1. Air Pollution in the environment

II. LITERATURE SURVEY

Clustering is one of the most widely used techniques in the analysis of time series data among all techniques applied. Objects with maximum similarities in a data set are grouped forming clusters [1]. A data-driven approach is used to detect environmental changes, based on direct and continuous sensor data and multi-sensor fusion [2]. The method of data fusion is used by extracting useful information from the raw data to reduce the data size and dimension. Multiple sensors support to improve the span of collecting data which enhances the depth of the results obtained from data fusion [3]. Common data fusion is the method of studying data value from various perspectives and transform the data into a single dataset with specific features [4]. Simulation-based and data mining- based methods are the two major methods used for predicting air pollution [5]. Pollution prediction is an important problem because pollution can affect individuals and their health. There are limitations when we use traditional air pollution prediction techniques [6]. From several sources, data or information is combined in a unified manner and Fuzzy logic is used to make decisions from the sensed data [7]. Machine learning techniques are used in air quality forecasting. A model can be efficiently trained by machine learning. For air pollution prediction, statistical models have been applied to the meteorological data [8]. To monitor the air pollution monitoring system is used. These air pollution monitoring systems are built using sensors. The polluted location is found using the sensors in which GPS is enabled [9]. The air quality index of a given area is predicted using machine learning. Using historical data, the model predicts the air quality index for a particular period in a particular area [10]. Machine learning algorithms are very efficient in the prediction of air pollution. Data from the sensors are gathered by artificial intelligence [11]. The pollution level in a particular area can be predicted and necessary actions can be taken by following that [12]. In this paper, a methodology measures the sum of energy about a linked node and it also takes the highest-energy route to transmit data to the surface location node. In case the surface position node is occupied with touch, the successor sub-aquatic neighbor will be given an update instruction to take a replacement surface position node to avoid data loss [13].

TABLE I
SENSORS USED TO MEASURE THE AIR POLLUTANTS

Sensors used	Used to measure	Usage	
Electrochemical Sensor	NO_2,SO_2 ,O_3,NO, CO	The chemical reaction of the air gasses to the electrode within a sensor in a liquid.	
Metal Oxide Sensor	NO_2,O_3, CO	Gases in the air react to the surface of the sensor and alter its resistance.	
Optical particulate counter	PM	Detect particulate emissions by measuring the particulate lights scattered.	
Optical Sensor	CO, CO_2	Gases such as carbon monoxide and carbon dioxide are measured by measuring the absorption of infrared radiation.	
Photoionization detector	VOC	Volatile organic compounds and the resulting electric current is measured.	

III. COLLECTION AND PREPROCESSING OF DATA USING DATA FUSION

A. Collection of Data

Data Fusion (DF) methods have gained increasing attention from several societies over the past few years due to the growing demand to incorporate the huge amount of data generated by the systems. Data Fusion is widely interpreted according to the applications, innovations, and relevant populations. In a broader context, it refers to the gathering and analysis of various data from multiple sensors to get the properties of an ecosystem or entity. DF provides marked benefits over data from a single source. Besides the statistical benefit acquired by integrating data from the same source,

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

making use of these multiple sensors data will improve the precision where a volume can be measured and categorized. In this paper, a real-time air pollution data are collected from 3 different air quality monitoring sites. The data were acquired from the Central Pollution Control Board, Chennai region. Collected the meteorological data which also, will influence the concentrations of the air pollutants, including relative humidity(RH), air temperature (Temp), Wind Direction (WD), Wind speed(WS) and Pressure, etc. The locations of three air quality monitoring stations are Alandur Bus Depot, Chennai. Latitude 12.9099161 and Longitude 80.1076538. Manali, Chennai Latitude 13.164544 and Longitude 80.26285. Velachery res area Latitude 13.0052189 and Longitude 80.2398125. Sensor fusion plays a role here. During the data collection, different sensors are used to collect the data, for example, particle sensor detects the physical contaminant. Humidity and temperature are detected using thermal sensors. Gases like O2, NO2, SO2, CO2, CO are monitored using gas sensors.□

B. *Preprocessing*

After collecting the data, Heterogenous data fusion is performed which means the different types of data from different types of sensors. Fuse all three different location data into one. However, at some hours the original data may have duplicate records or missing incomplete values. Because of the huge amount of fluctuations and situations involved, which produce noise in timelines, the collected time series had to be revised and renewed. So as per the fusion guidelines, preprocessed data must be clustered or classified then filter the data and make a decision based on that. So here we proposed an LDA for clustering. After a classification Random forest tree is proposed for decision making.

Here we are analyzing the Pollution level for three different locations, hence here we do not require a particular date or time.

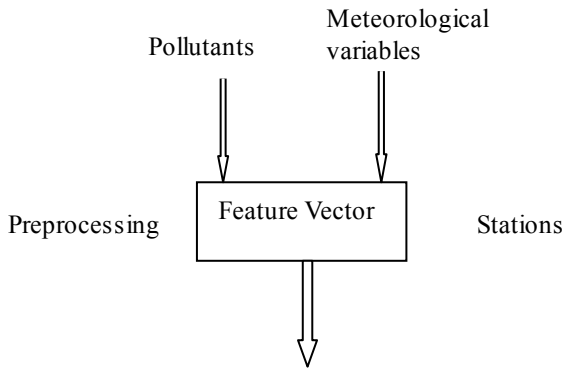


Fig. 2. Preprocessing for all the three stations

Proposed a system for air quality monitoring and tracking of the Chennai. Meteorological variables At the data fusion stage wind speed and wind direction are used as a deciding factor. Fig 3,4,5,6,7and 8 shows particle matters with different gases and meteorological variables. A data fusion framework provides many benefits over individual sensors, for instance, improved measurement through multiple measurements of a physical phenomenon when it comes to the simple task of using and transmitting data to a specific task.

IV. FEATURE SELECTION TECHNIQUE

A. *Linear Discriminant Analysis(LDA):*

LDA is a dimensionality reduction technique. It is used to reduce the dimensions. It is used to find differences in groups and partition two or more classes. LDA is one of the clustering techniques which is used to group the data and also for partitioning data and gives the best accuracy than compared to other algorithms. Apply LDA over the equivalent dataset twice with various jobs each time. In the primary methodology, LDA will fill in as a classifier and posterity it will decrease the dimensionality of the dataset and a neural system will play out the grouping task, the consequences of the two methodologies will be analyzed a short time later.

LDA makes a few rearrangements about the information:

1. That your information is Gaussian, that when plotted, every element is molded like a chime bend.
2. That there is a similar shift each time, that estimates of each element fluctuate across the mean by a similar amount of all items considered.

The LDA model gages the mean for each class with these presumptions, and adjustments from your knowledge. In the case of two groups in the univariate (single knowledge variable) it is anything but hard to understand this. In the traditional direction, the mean (m) estimate of each information (xi) for each class (ki) can be calculated, by dividing aggregate qualities by the absolute number of qualities.

$m = 1/nki * sum(xi)$

For class k, where m is the mean estimate of x, nki is the sum of occasions with class k. The adjustment is defined as the standard squared differentiation of each reward from the mean for all groups.

$sig^2 = 1 / (n-Ki) * sum((xi– m)^2)$

Where sig^2 is the variance of all inputs (xi), n is the number of instances, Ki is the number of groups, and m is the mean of xi.

B. *Predictions made with LDA:*

LDA does the prediction process by estimating the likelihood that it will be an input set belonging to each group. The group with the highest amount of probability is the class of output, and the process of prediction is made. The model here makes use of the Bayes Theorem to estimate probabilities. It is used to calculate the probabilities of the

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

given output class (ki), the input (xi) using each class's probability, and the likelihood of each class's data.

P(Y=x|iX=xi) = (Pok * fki(xi)) / sum(PIl * fl(xi))

Where Pok refers to the basic likelihood of each class (ki) found in the training. In Bayes' theorem, this is called the prior probability

Pok = nki/n

The above f(xi) is the approximate likelihood of xi belonging to a given class. For f(xi) a Gaussian distribution function is used. By solving the above equation, we get

D(xi) = xi * (m/sig^2) – (m^2/(2*sig^2)) + ln(Pok)

D(xi) is the discriminating function for class ki given with input xi, from which all them, sig2, and Pok are obtained.

Firstly, all the characteristics of that dataset that your work will be guiding you about the decision of applying LDA is a linear classifier or algorithm for the reduction of dimensionality to perform classification. The prior aim of, In fact, Linear Discriminant Analysis is an example of classes that slowly transfer them into a specific feature space, if the dataset is linearly separable, applying LDA as a classifier is the only way to achieve better results. Next, if the dataset is no longer linear separable, the LDA can seek as much as possible to arrange the data set in another space as the optimal linear separability, although it may still be an example of class duplication due to non-linear data characteristics. In this case, you have to need to use another classification technique to deal with nonlinear data, such as a neural network model with layers that are hidden, a neural network with radial basis functions that are available.

C. *Random Forest*

It is one of the managed arrangement calculation. It orders the huge dataset which contains enormous size. It is a choice calculation that is utilized for the group to give a precis ˆ cast. The irregular timberland calculation is a mana¡ arning model; it utilizes marked information to "realize" how to arrange unlabeled information. This is something contrary to the K-implies Cluster calculation. The Random Forest Algorithm can be used to resolve relapse issues as well as structural problems, making it an assorted model that engineers usually use. While performing the Random Forest algorithm, based on data classification, you must be informed that you are using the Gini index or the algorithm used to validate how the decision tree branch nodes are used.

$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2$$

This method uses the class and probability to calculate Gini on a node for each branch, which branches are more likely to occur. Here pi is the class frequency you see in the dataset and c shows the number of classes. Entropy can also be used to determine how nodes in a decision tree can branch out.

$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i)$$

A single decision tree must be understood before we truly understand the random forest. We need to consider the variations between the present node, branch, and leaf and how different formulae can be used to finish the decision-making process. The random forest algorithm would be useful for all types of data sets, whether it's regression or classification data if performed correctly. It is simple to use, fast to train and gives an accurate image of the decision trees that it uses.□

V. RESULTS AND DISCUSSION

Real-time data is gathered from a different place and different time. We performed data analysis and gives the prediction which gives much accuracy. After analyzing the data from the datasets, the locations are identified as dangerous ones based on most severity and weather conditions. Fig 3-7 gives the various gas level based on PM2.5.

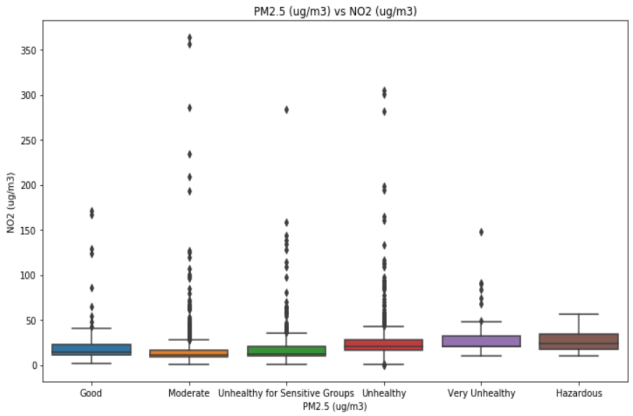


Fig. 3. PM2.5 vs NO2

In fig.3 PM2.5vs NO2 is shown. NO2 affects the lungs and enhances breathing infections which may lead to asthma problems.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

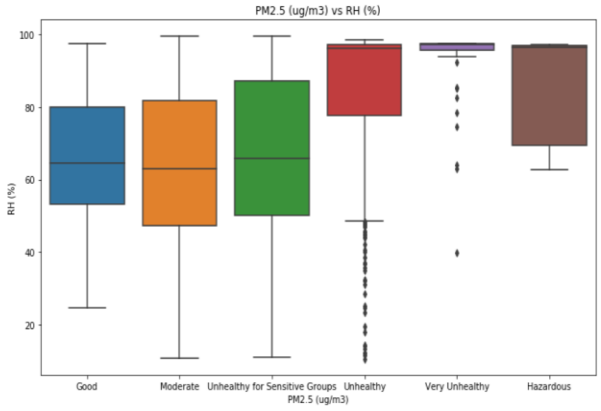


Fig. 4. PM2.5 vs RH

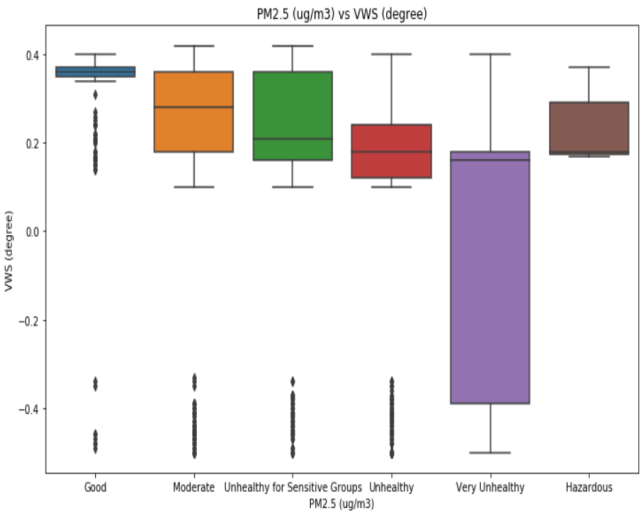


Fig. 5. PM2.5 vs VWS

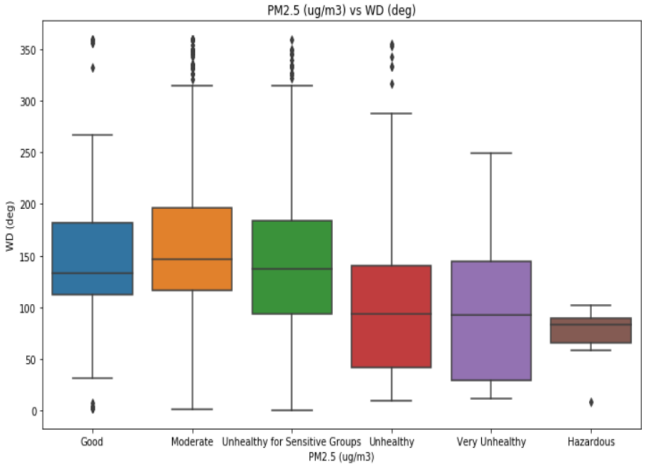


Fig. 6. PM2.5 vs WD

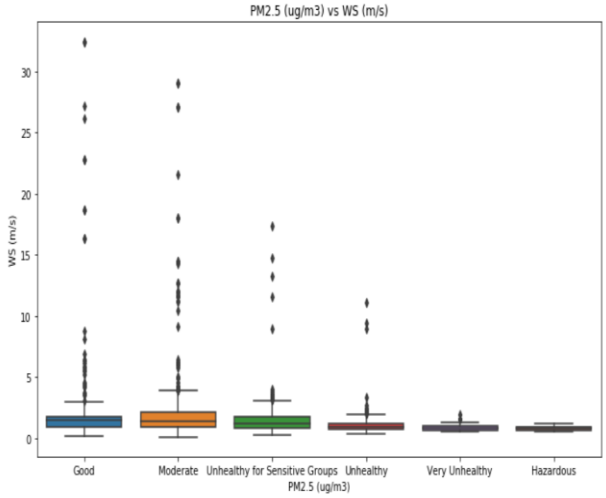


Fig. 7. PM2.5 vs WS

The direction of the wind is an important parameter that affects PM2.5. The wind carried an equal amount of contaminants from different directions. Fig5-7 gives the PM2.5 concentration level based on wind speed, wind direction, and relative humidity.

VI. CONCLUSION

The proposed work can predict the gas level by combining multiple sensor data from multiple cities and predict the pollution level for different gases using LDA and Random forest tree. We have collected data from the Central Pollution Control Board. Analyzed the data using python.

ACKNOWLEDGMENT

We are thankful to the Sathyabama Institute of Science and Technology management for providing all the requisite facilities to prepare this manuscript.

REFERENCES

[1] Yunliang Chen, Lizhe Wang , Fangyuan Li , Bo Du , Kim-Kwang Raymond Choo ,Houcine Hassan, Wenjian Qin, Air quality data clustering using EPLS method R, Information Fusion(2016)

[2] Eiman Kanjo a, Eman M.G. Younis b, Nasser Sherkat a, Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach, Information Fusion 40(2018) pp 18-31

[3] Wenxiu Ding a , Xuyang Jing a , Zheng Yan , Laurence T. Yang c , A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion, Information Fusion 51 (2019) 129–144

[4] Jia Liu, Tianrui Li, Peng Xie, Shengdong Du, Fei Teng , Xin Yang, Urban big data fusion based on deep learning: An overview, Information Fusion 53 (2020) 123–133

[5] Hamed Karimian, Qi Li, Chunlin Wu, Yanlin Qi, Yuqin Mo, Gong Chen, Xianfeng Zhang, Sonali Sachdeva, Evaluation of Different Machine Learning Approaches to Forecasting PM 2.5 Mass Concentrations, *Aerosol and Air Quality Research*, 19: 1400–1410, 2019

[6] R. O. Sinnott and Z. Guan, "Prediction of Air Pollution through Machine Learning Approaches on the Cloud," 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT), Zurich, 2018, pp. 51-60.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

[7] Air Pollution Detection Using Multisensor Data Fusion , Md Faruck , Punyasha Chatterjee, Workshops ICDCN '18: Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and NetworkingJanuary 2018 Article No.: 24 Pages 1–2

[8] Dixian Zhu, Changjie Cai, Tianbao Yang and Xun Zhou, A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization, Big Data Cogn. Comput. 2018, 2, 5; doi:10.3390/bdcc2010005

[9] A.Sivasangari, P.Ajitha, K.Indira, Air Pollution Monitoring and Prediction using Multi View Hybrid Model, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 6S4, April 2019

[10] Mrs. A. Gnana Soundari MTech, (PhD) , Mrs. J. Gnana Jeslin M.E, (PhD) , Akshaya A.C, INDIAN AIR QUALITY PREDICTION AND ANALYSIS USING MACHINE LEARNING, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 11, 2019

[11] Kostandina Veljanovska, Angel Dimoski , Air Quality Index Prediction Using Simple Machine Learning Algorithms International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 7, Issue 1, January - February 2018

[12] Aditya C R, Chandana R Deshmukh , Nayana D K , Praveen Gandhi Vidyavastu, Detection and Prediction of Air Pollution using Machine Learning Models, International Journal of Engineering Trends and Technology (IJETT) – volume 59 Issue 4 – May 2018

[13] R.M.Gomathi, J.Martin Leo MAnickam, "Energy Efficient Shortest Path Routing Protocol for Underwater Acoustic Wireless Sensor Network ",Wireless Personal Communications, 98(1), pp. 843–856, Jan 2018.

[14] J.M. Barron-Adame, M.G. Cortina-Januchs and A. Vega-Corona," Data Fusion and Neural Network Combination Method for Air Pollution Level Monitoring", 978-1-4244-3760-3/09/$25.00 c 2009 IEEE, 2009 7th IEEE International Conference on Industrial Informatics (INDIN 2009)

[15] Andrew Fox, B. V. K. Vijaya Kumar, Jinzhu Chen, Fan Bai "Multi-Lane Pothole Detection from Crowdsourced Undersampled Vehicle Sensor Data" IEEE Transactions on Mobile Computing, Vol. 16, Issue. 12, 2017, 3417-3430.

[16] Jeong-Joo Kim, Su-il Choi, "Implementation of Pothole Detection System Using 2D LiDAR", International Conference on Electronics, Information, and Communication,2017, 607-610.

[17] https://ec.europa.eu/environment/air/pdf/Brochure%20lowercost%20sensors.pdf

[18] [2] Andrew Fox, B. V. K. Vijaya Kumar, Jinzhu Chen, Fan Bai "Multi-Lane Pothole Detection from Crowdsourced Undersampled Vehicle Sensor Data" IEEE Transactions on Mobile Computing, Vol. 16, Issue. 12, 2017, 3417-3430.

[19] Jeong-Joo Kim, Su-il Choi, "Implementation of Pothole Detection System Using 2D LiDAR", International Conference on Electronics, Information, and Communication, 2017, 607-610.

[20] Wang, Jianhua & Ogawa, Susumu. (2015). Effects of Meteorological Conditions on PM2.5 Concentrations in Nagasaki, Japan. International journal of environmental research and public health. 12. 9089-101. 10.3390/ijerph120809089.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Fifth International Conference on Communication and Electronics Systems (ICCES 2020)
IEEE Conference Record # 48766; IEEE Xplore ISBN: 978-1-7281-5371-1

An Efficient Model for Predicting Brain Tumor using Deep Learning Techniques

Meenakshi. A
U.G Scholar
Department of Information Technology
Sathyabama Institute of Science and Technology
India.

S. Revathy
Associate Professor
School of Computing
Sathyabama Institute of Science and Technology
India.

***Abstract* — One of the most popular applications of Artificial Intelligence that has seen an immense growth in the digital era is Deep Learning Techniques, where the system studies and improves its performance through progressive learning without any explicit programming. Deep learning is widely used in numerous applications one of them being medical analysis. Feature extraction and Image classification are considered to be the most popularly used approaches done using deep learning process. Prior to this, some neural networks were used for image Classification. A significant, practical deep learning model is proposed which uses back propagation neural network classifier to predict brain stroke through CT/MRI scan images. The performance and accuracy of the proposed model is evaluated and compared with existing models and it produces high sensitivity, specificity, precision and accuracy.**

Keywords — ***Brain tumor, Classification, Deep Learning, Fusion, Medical Image Analysis.***

I. INTRODUCTION

Brain tumor is an irregular tissue growth in the brain. Humans are significantly affected by brain tumor due to the irregular growth of cells within the brain that can impair brain activity and functions and can also lead to other life threats Malignant tumors and
benign tumors are the significant class of brain
tumors found out. MRI- Medical imaging technique is used to gain visual image of the interior of the human body and this approach can be used to diagnose non-invasive ability in the human body. Ultrasound, X-ray, single photon emission computed tomography, computed tomography, positron emission topography and MRI are some of the techniques which are based on non-invasive approaches. Magnetic Resonance Imaging (MRI) provides comparison images of the brain and Cancer tissues as well. This is why MRI is often used when compared with other medical imaging techniques. Many machine learning models were built using computational data set as well as images to predict brain tumors. The proposed work uses a deep learning model that uses neural function for predicting brain tumor through reverse propagation.

II. RELATED WORK

Number of research works were provided by various researcher scholars for segmenting or classifying a specific object or a certain area from an image. Deep learning is one of the most commonly used classification techniques used by numerous researchers for classification of the tumor cells.
One of the most vital part of deep learning approach is the extraction of features from the dataset that is being given and trained [1] Images pertaining to the measurement of various parts of the human body based on different scales like microscopic or macroscopic are coined as biomedical images. These biomedical images are generated through various instruments such as Ultrasound machines and CT scan images [2]. Examining the capability of supervised machine learning classifiers in extracting flooded areas [3].An assessment of support vector machines for land cover classification [4]. A review of classification techniques was performed by Zaharakis [5]. Unsupervised feature learning for

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

aerial sense classification was performed by cheriyadat [6]. Fully automatic segmentation using support vector machine classification [7]. Fusing decision trees based on genetic programming for classification [8]. Naïve Bayesian classification for predicting the recurrent events in breast cancer were performed by Dumitru [9].A brain tumor was detected by Jason J. Cors [10] by integrating segmentation with Bayesian model.Weighted aggregation algorithm was used for the detection of tumor cells. The performance evaluation was done on a larger dataset where stochastic models were used extracting the features. The model could be enhanced by providing an accurate boundary of the tumor cells. Concurrent image segmentation and bias correction were performed by Kaihua Zhang [11], where minimization of energy was performed by an efficient Bayesian Learning Approach. The technique stated that the experimental results were performed on a real dataset of images and produced an accurate intensity of homogeneity. Md. Badrul et al., [12] classified lung cancer images from CT scan using MLFFNN technique yielding a good accuracy of 96.67%. FFNN approach used in [13], described the classification technique producing an accuracy of 92%. Rajesh Kumar Tripathy [14] combined SVM with Least Square and provided an accuracy of 95.34%. Persi et al [15] used the Particle Swarm Optimization technique for predicting heart disease yielding 92.2% accuracy. Bashar A [16], did a survey on evolving deep learning neural network architectures. Bindhu V [17],described biomedical image analysis using semantic segmentation. C. Kalpana et al [18] ,did a survey on policy combined and time bound hierarchical delegation model in electronic health records. Kalpana, C., et al [19] , performed an Analysis of time bound collaborative access control delegation model in electronic health records.

III. PROPOSED WORK

Fig. 1 represents the steps for image preprocessing. The enhanced and efficient model designed in the current paper first extracts the sequence of images from the scanner database. The images are then pre-processed where all the features of the images are extracted during neural networks. The preprocessing is initially done by Medical Image Analysis where all the images are processed in such a way that it generates new images that have high quality when compared to the original images. The Input impression is provided as a result of the scanner repository and it is initially changed into Grayscale. The Discrete Cosine Transform is used towards the picture whenever the distinctions on the picture is improved in such a manner that the little options that come along with the pictures are completely apparent. The input image is broken down into 8-by-8 or 16-by-16 blocks, and the two-dimensional DCT is determined for each block.The DCT coefficients are then quantized, labelled, and forwarded. The JPEG receiver (or JPEG file reader) decodes the quantized DCT coefficients, calculates each block's inverse two-dimensional DCT and then reassembles the blocks into a single image For typical images, many of the DCT coefficients have values close to zero. The N-by-M matrix A's two-dimensional DCT is defined as follows in equation 1:

$$F(u,v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i).\Lambda(j).cos\left[\frac{\pi.u}{2.N}(2i+1)\right] cos\left[\frac{\pi.v}{2.M}(2j+1)\right].f(i,j)$$

Equation 1

Where,

M = row of the image matrix

N = column of the image matrix

Δ(i) = data sequence of length i

Δ(j) = data sequence of length j

f(i,,j) is the pixel intensity at row I and column j.

F(U,V) = DCT coefficient in row k1 and in column k2.

The input DCT is an 8 by 8 integer number. This array contains the gray scale amount for every pixel.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

Numerous feature extraction methods and classification techniques are used widely to identify and detect the location of the tumor cells in a human body. In this paper, deep learning technique known as Backpropagation is focused majorly. Backpropagation is a type of artificial neural network which comprises of three layers, of units namely input, hidden and output. The various steps involved in the entire model is discussed as follows: CT/MRI images obtained from patients suffering from brain stroke are collected from Database and hospitals. The file formats collected from the database are of .jpeg, .tift, .png file formats and the file dimensions consist of rows and columns. Subsequent to the image acquisition stage, the images pass through the preprocessing stage. Feature Extraction is a technique of recording visible written content of all the pictures. Functions that are obtained from the segmented picture include spot, mean, irregularity index, equivalent diameter, perimeter, regular deviation as well as entropy. Extracted Features are utilized inside a neural classifier to instruct the product inside such a manner which it might acknowledge a specific category as abnormal or normal. The classifier is going to assign the unfamiliar item on the appropriate category according to the extracted characteristics.

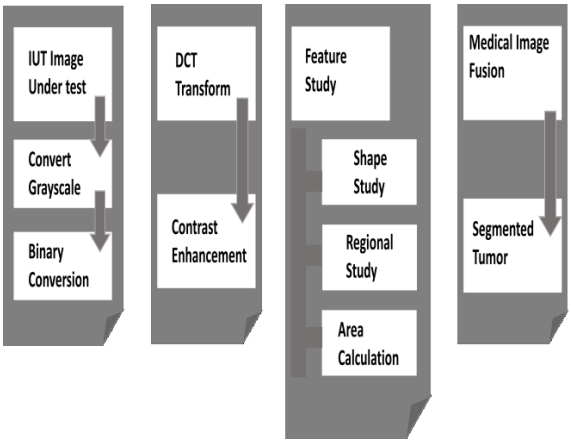


Fig.1 Image Preprocessing System Architecture

In segmentation process, the homogeneous regions are obtained from the input image. The interested regions of an image are found using the segmentation process. It reduces the number of pixels of an image to make it easier for the next step of feature extraction and classification. Segmentation is tougher in CT scan images and when combined with a huge amount of data because of the extra dimensions that need to be considered for the neighborhood calculation.

Classification using Backpropagation is a type of artificial neural network consisting of three layers of units,namely input ,hidden and output. The input unit is the raw information sent to the network.The performance of each hidden unit is determined by the input units working and the weights between the hidden and the output units.The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units. Neural networks are a series of interconnected nodes, called neurons or perceptrons. Growing neuron takes one piece of input data, usually one pixel of the image, and applies a simple calculation, called a result-generating activation function. Each neuron has a numerical weight which affects their outcome. This results in additional neural layers before the cycle is completed. This result is fed into additional neural layers before the neural network produces a prediction for each input or pixel, repeated for a large number of images at the end of the process, and the network learns the most suitable weights for each neuron in a process called backpropagation. The backpropagation algorithm works by calculating the gradient of the loss function by chain rule with respect to each weight, calculating the gradient one layer at a time, iterating backwards from the last layer to prevent redundant measurements of intermediate terms in chain rule.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

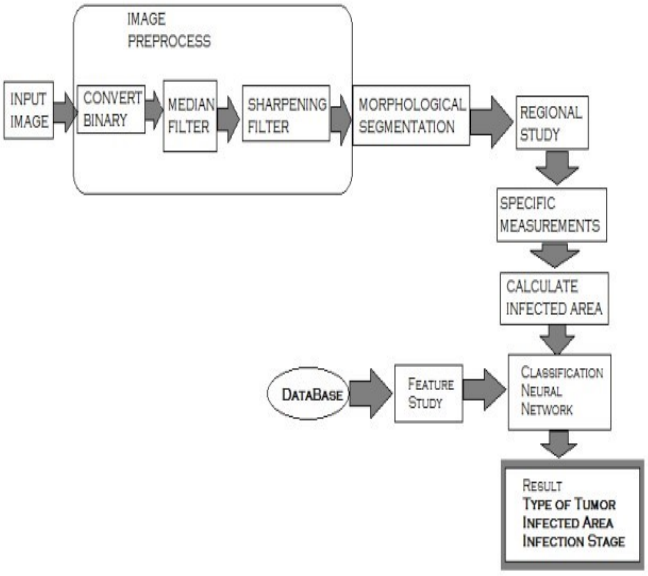


Fig. 2 Workflow of Proposed work

Fig. 2 depicts the flow chart of the entire classification process.

PROPOSED ALGORITHM:

Step 1: Collect the CT/MRI images from hospitals.

Step 2: Preprocess the collected images using Medical image analysis.

Step 3: Segmenting the tumor using bounding box.

Step 4: Classifying the input images using Backpropagation.

Step 5: Using the classified images predict whether tumor is there or not.

IV. EXPERIMENTAL RESULTS

The tests are carried out on MATLAB R2017b model. The calculations are carried out by using picture Classification Learner Toolbox which is conveniently obtainable in MATLAB. First, the type in picture via the scanner repository is provided to the device. Fig. 3, is a sample input image obtained from a scanner database. When this image is passed the pre-processing takes place where all the features are extracted for further segmentation.

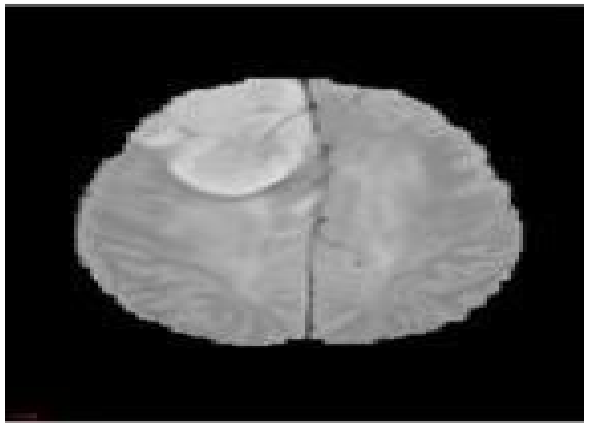


Fig. 3: Input Image

As the next step, the RGB components are extracted from the image and then converted to Grayscale image. Fig.3 shows the output image when the Grayscale filter is applied to the image.

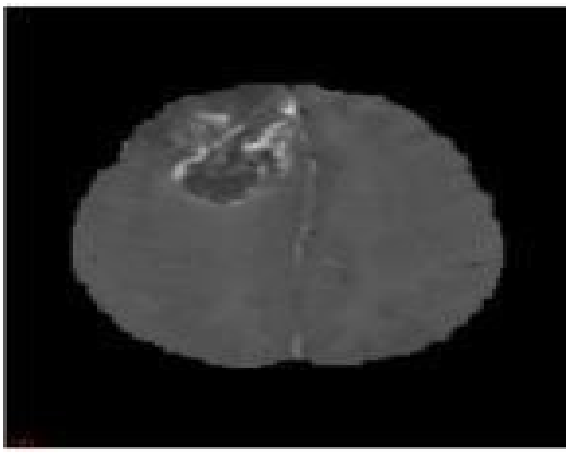


Fig. 4 FilteredImage

Both Fig. 3 and Fig. 4 are compared with each other to find the difference between the features and then Medical image fusion is applied to the images for tumor detection

The features are then trained with Deep Neural Networks where some images are kept for training and the rest are kept for testing of images.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

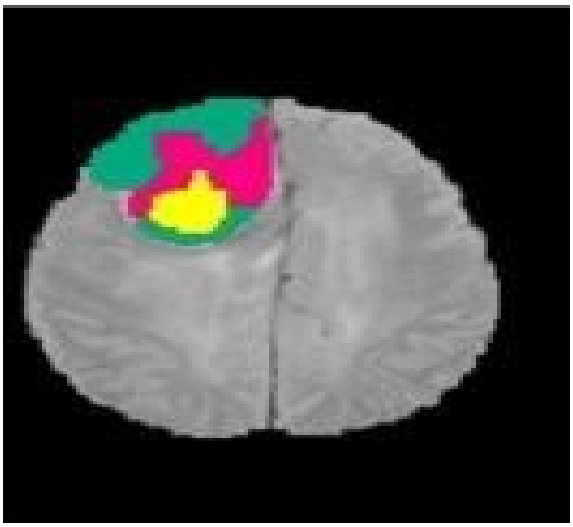


Fig. 5 Image Segmentation

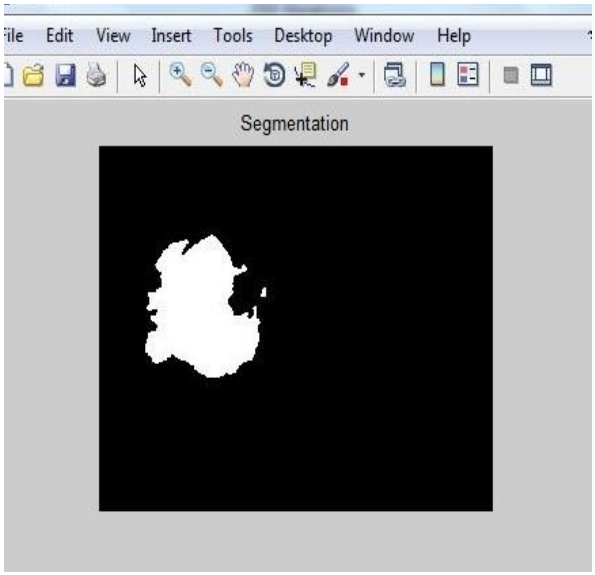


Fig. 6 Segmented Image

The classifier learns the features and the successfully classifies when any new image is given to it. It identifies where the image is benign or malignant. Fig. 5 shows the identification of the tumor cell from the image and Fig.6 show the segmented tumor from the image. The accuracy of the network is observed with 99.69% whereas 99.9% specificity and 80.48% sensitivity. The segmentation of the tumor from the exact image gives us various parameters for measuring such as its intensity, volume, and size. This helps in diagnosing and treating the tumor cells more effectively.

Fig 7.a Input Preprocessed images for Classification

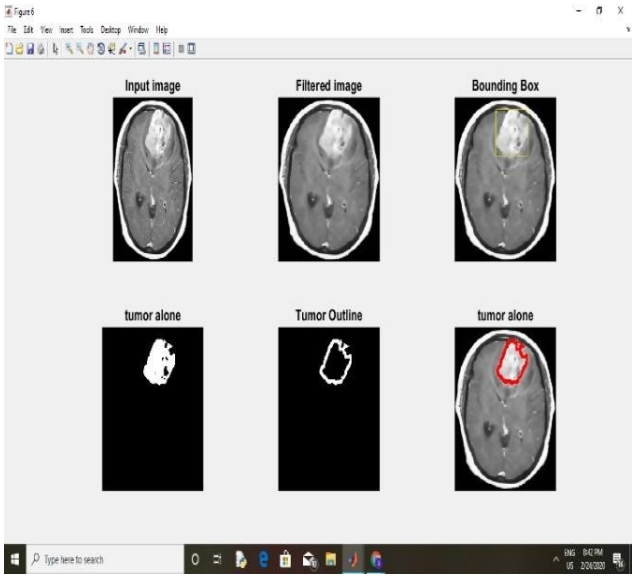
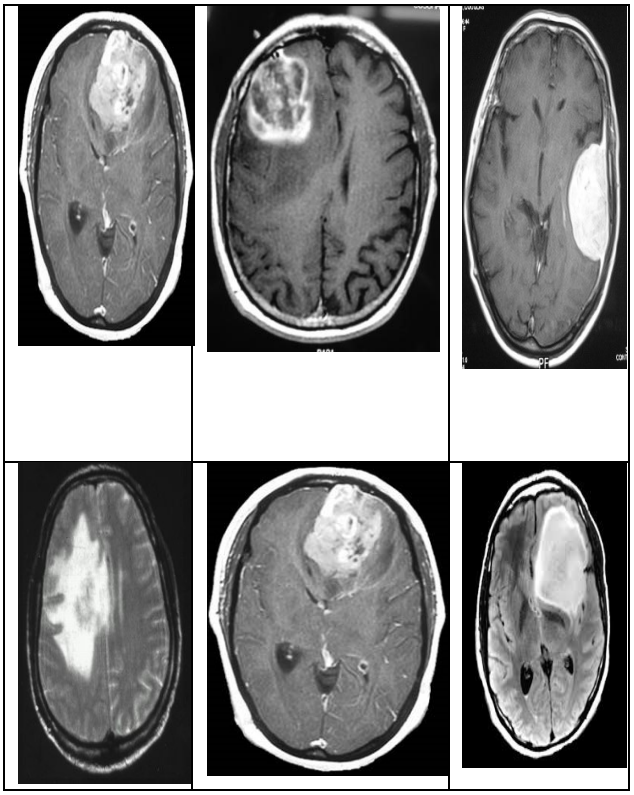


Fig. 7.b Input preprocessed images for Classification

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.



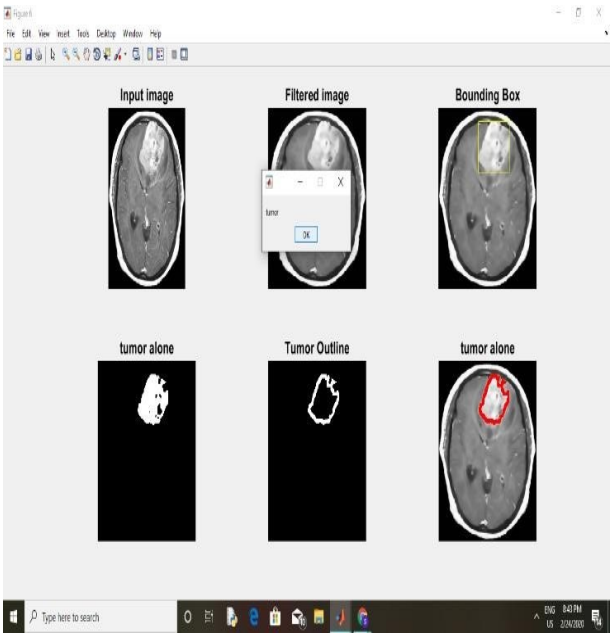


Fig. 8 Tumor Prediction using classification

In Fig. 7.a &b, the input preprocessed images for classification in fig, 8, the detection message box when this system was achieved as shown in Fig. 9. Finally this system was completed through the deep neural networks by MATLAB.

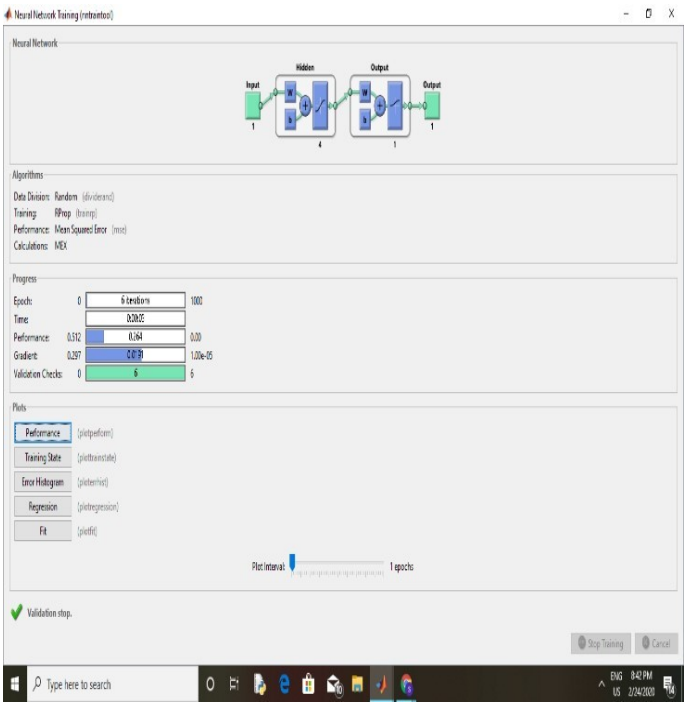


Fig. 9 Neural network Classifier.

V. Performance Analysis

The accuracy,Sensitivity,Specificity and precision of proposed work is compared with existing machine learning models and it is given in has table 1.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

Table1: Comparison of proposed work with existing work

Parameters	Probabilistic neural network	Autoencoder	Backpropagation (proposed)
Accuracy	83%	85%	92%
Sensitivity	48.02	52.03	60.05
Specificity	39.50	41.05	46.76
Precision	42.00	46.06	57.02

VI. CONCLUSION

Brain Tumor is an irregular brain tissue growth that affects normal brain functions. The main goal of medical image processing is to classify correct and relevant details with the least possible error using the images. Identification of brain tumors via MRI images is a difficult task because of the structure of the brain. Recognition of brain tumors using MRI scans is not especially difficult due to brain sensitivity. So in the proposed work, deep learning model has been constructed to predict brain tumor with high accuracy rate and low time duration. Brain tumors can be detected by means of detailed MRI images that can be categorized into four different groups, including pre-processing, image segmentation, feature extraction, and image grading. There is no way to prevent Brain Tumor , but there are treatment options .Early Segmentation and Analysis is most helpful in predicting whether tumor is there or not. However, now prediction is made using MRI images. In future, a fusion of CT and MRI images can be used to classify and predict Tumor.

Acknowledgement

I thank Nandini Hospital for guiding and supporting me in creating the dataset.

REFERENCES

1. Rahman, M. M., Bhattacharya, P., & Desai, B. C. (2007). A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE Transactions on Information Technology in Biomedicine, 11(1), 58-69.
2. Morales, M., Tapia, L., Pearce, R., Rodriguez, S., & Amato, N. M. (2004). A machine learning approach for feature-sensitive motion planning. In Algorithmic Foundations of Robotics VI (pp. 361-376).
3. Berlin, Heidelberg(2015). Examining the capability of supervised machine learning classifiers in extracting flooded areas from Landsat TM imagery: a case study from a Mediterranean flood. Remote sensing, 7(3), 3372-3399.
4. Huang, C., Davis, L. S., & Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. International Journal of remote sensing, 23(4), 725-749.
5. Kotsiantis, S. B., Zaharakis, I., &Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-24.
6. Cheriyadat, A. M. (2014). Unsupervised feature learning for aerial scene classification. IEEE Transactions on Geoscience and Remote Sensing, 52(1), 439-451.
7. Bauer, S., Nolte, L. P., & Reyes, M. (2011, September). Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 354-361). Springer, Berlin, Heidelberg.
8. Liu, K., Tong, M., Xie, S., & Zeng, Z. (2014, August). Fusing decision trees based on genetic programming for classification of microarray datasets. In International Conference on Intelligent Computing (pp. 126-134). Springer, Cham.
9. Dumitru, D. (2009). Prediction of recurrent events in breast cancer using the Naive Bayesian classification. Annals of the University of Craiova Mathematics and Computer Science Series, 36(2), 9296.
10. Corso, J. J., Sharon, E., Dube, S., ElSaden, S., Sinha, U., & Yuille, A. (2008). Efficient multilevel brain tumor segmentation with integrated bayesian model classification. IEEE transactions on medical imaging, 27(5), 629-640.
11. Zhang, K., Liu, Q., Song, H., & Li, X. (2015). A variational approach to simultaneous image segmentation and bias correction. IEEE Transactions on Cybernetics, 45(8), 1426-1437.
12. Md. BadrulAlam Miah, Mohammad Abu Tousuf, Detection of Lung Cancer from CT Image Using Image Processing and Neural Network, IEEE, In Proceedings of 2nd Int'l Conference on Electrical Engineering and Information & Communication Technology, 2015
13. Shubhangi Khobragade, AdityaTiwari, C.Y.Patil , VikramNarke,Automatic Detection of Major Lung Diseases using Chest radiographs and Classification by Feed-forward Artificial Neural Networks, In Proceedings of 1stIEEEInternational Conference on Power Electronics, Intelligent Control and Energy Systems, pp.1-5, 2016.
14. Rajesh Kumar Tripathy, SailendraMahanta Subhankar Paul, Artificial Intelligence – based Classification of Breast Cancer using Cellular images RSC Advances , Issue 18, Page 8939 to 9411, 2014 .
15. Persi Pamela I, Gayathri P, N. Jaisankar, Fuzzy optimization Techniques for the Prediction of Coronary Heart Disease Using Decision Tree, International Journal of Engineering and Technology, Vol 5,No 3,pp. 2506-2513, 2013.
16. Bashar, A. (2019). SURVEY ON EVOLVING DEEP LEARNING NEURAL NETWORK ARCHITECTURES. Journal of Artificial Intelligence, 1(02), 73-82.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

17. Bindhu, V. "Biomedical Image Analysis Using Semantic Segmentation". Journal of Innovative Image Processing (JIIP) 1, no. 02 (2019):91-101.

18. C. Kalpana, et al A survey on policy combined and time bound hierarchical delegation model in electronic health records,Jour of AdvResearch in Dynamical & Control Systems, Vol. 10, 03-Special Issue, 853- 855,2018.

19. Kalpana, C.,et al "Analysis of time bound collaborative access control delegation model in electronic health records. In 2017 International Conference on Computing Methodologies and Communication (ICCMC), pp. 424-429. IEEE, 2017.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 15:09:36 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020)
IEEE Xplore Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2

Air Pollution Detection and Prediction Using Multi Sensor Data Fusion

E. Brumancia, S. Justin Samuel, L.Mary Gladence
easpinbrumancia@gmail.com

***Abstract*— The level of air pollution is rising day by day which affects human life and the environment. Hence, it is important to control it by keeping an aircheck at its level constantly. To reduce the pollution level we develop a multi-sensor data fusion model that detects and predicts the most severe gas. In this work we suggest an effective algorithm to cluster the multiple sensors data which used to group the data and partition the data. This paper proposes a random forest approach to predict air quality. The efficiency is measured with real multiple city-data.**

***Index Terms*— Multisensor, data fusion, air pollution, prediction, Linear Discriminant Analysis, Random Forest**

I. INTRODUCTION

Air pollution is a combination of organic and man-made pollutants in the air we breathe. It is typically classified into two divisions: external air pollution, and internal air pollution. External air pollution entails exposures to un-built ecosystems. For example, Pollution from such sources as cars and power stations, factories, etc are polluting the air. Whereas indoor air pollution is from household products, chemicals, building materials, tobacco smoke, gases, etc. The air quality control system has been used to monitor air pollution levels. Over the past 30 years, studies have discovered a wide range of health consequences that are thought to be linked with air pollution exposure. Some air pollutants are harmful. Inhalation of these pollutants can raise a chance of having problems with your health. People with lung or heart problems, Children's and late teenagers are at a greater risk of emissions from the atmosphere. It doesn't mean we have pollution only outside, the air inside buildings can also be toxic and harm your health. Air pollution is defined as the presence of chemical components known as pollutants in the air. The release of chemicals into the atmosphere leads to air pollution. These pollutants change the composition of the air, making it hazardous. The main source of air pollution is fossil-fuel burning as fossil fuels are burned a lot of chemical components are released in the air which causes air pollution. Air pollution has its implications on humans and the environment. This causes human respiratory and heart problems. When the polluted air is inhaled by humans the organs are affected. The other effects are global warming, industrial pollution, eutrophication, ozone depletion, etc. Air pollution can be reduced by using public transport, using energy-efficient devices, using solar energy and wind energy, etc. Using public transport instead of private transport reduces the emission of pollutants into the atmosphere. Using energy-efficient devices and renewable resources like solar power geothermal power can lower the consumption of electricity produced by burned fossil fuels.

Poor air quality is killing the people. The World Health Organization states Poor outdoor weather in 2016 caused an estimated 4.2 million premature deaths, In low- and middle-income nations, about 90 percent. Three billion citizens who are cooking or heating their homes by burning wood, coal, and gasoline get health problems due to indoor smoke. Higher rates of air pollution have been linked with a tumor, cardiac disease, stroke, and breathing disorders like asthma. Around 134 million Americans, according to figures by the American Lung Association. Owing to air pollution and premature death, about 40 percent of the population is at risk of disease. Major air pollutants are particulate matter, lead, sulfur dioxide, asbestos, carbon monoxide. Particulate matter smaller than 10 micrometers poses a higher risk to health, breathing deep into the lungs, then moving through the bloodstream.

The integration of multiple data sources to produce more consistent data is known as data fusion. Sensor fusion is defined as the combination of multiple sensors data to provide descriptions about an environment. Sensor fusion which is a subset of information fusion is also called Multisensor fusion. The multi-sensor data fusion can be further classified into homogeneous data fusion and heterogeneous data fusion. Homogeneous data fusion is the fusion of data of the same type of sensors and heterogeneous fusion of data integrates multiple sensor data. The different levels of data fusion processing are a fusion of the signal level, a fusion of the object level, a fusion of the function level, and fusion of the decision level.

Air pollution sensors are devices used when measuring air pollution. Those sensors track and detect pollutant concentrations in the air. These sensors majorly focus on five components, they are ozone, particulate matter, sulfur dioxide, carbon monoxide, and nitrous oxide. These are very costly sensors when released, but with technological growth, they have become affordable now.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

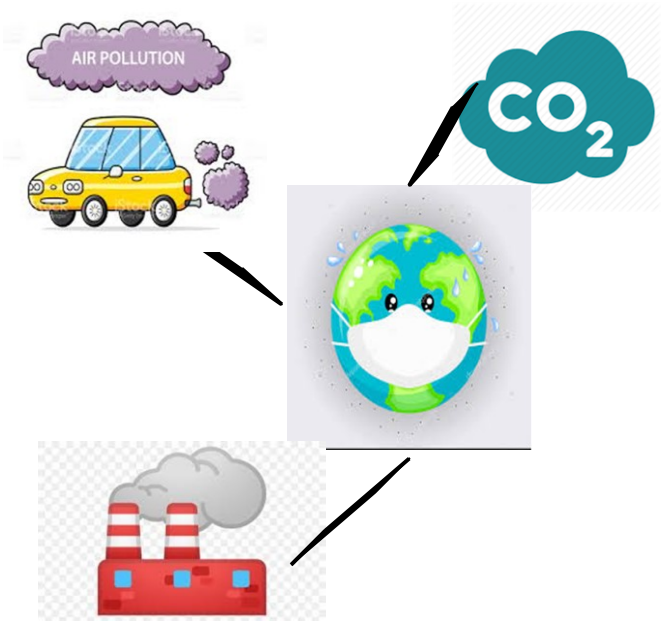


Fig. 1. Air Pollution in the environment

II. LITERATURE SURVEY

Clustering is one of the most widely used techniques in the analysis of time series data among all techniques applied. Objects with maximum similarities in a data set are grouped forming clusters [1]. A data-driven approach is used to detect environmental changes, based on direct and continuous sensor data and multi-sensor fusion [2]. The method of data fusion is used by extracting useful information from the raw data to reduce the data size and dimension. Multiple sensors support to improve the span of collecting data which enhances the depth of the results obtained from data fusion [3]. Common data fusion is the method of studying data value from various perspectives and transform the data into a single dataset with specific features [4]. Simulation-based and data mining- based methods are the two major methods used for predicting air pollution [5]. Pollution prediction is an important problem because pollution can affect individuals and their health. There are limitations when we use traditional air pollution prediction techniques [6]. From several sources, data or information is combined in a unified manner and Fuzzy logic is used to make decisions from the sensed data [7]. Machine learning techniques are used in air quality forecasting. A model can be efficiently trained by machine learning. For air pollution prediction, statistical models have been applied to the meteorological data [8]. To monitor the air pollution monitoring system is used. These air pollution monitoring systems are built using sensors. The polluted location is found using the sensors in which GPS is enabled [9]. The air quality index of a given area is predicted using machine learning. Using historical data, the model predicts the air quality index for a particular period in a particular area [10]. Machine learning algorithms are very efficient in the prediction of air pollution. Data from the sensors are gathered by artificial intelligence [11]. The pollution level in a particular area can be predicted and necessary actions can be taken by following that [12]. In this paper, a methodology measures the sum of energy about a linked node and it also takes the highest-energy route to transmit data to the surface location node. In case the surface position node is occupied with touch, the successor sub-aquatic neighbor will be given an update instruction to take a replacement surface position node to avoid data loss [13].

TABLE I
SENSORS USED TO MEASURE THE AIR POLLUTANTS

Sensors used	Used to measure	Usage	
Electrochemical Sensor	NO_2,SO_2 ,O_3,NO, CO	The chemical reaction of the air gasses to the electrode within a sensor in a liquid.	
Metal Oxide Sensor	NO_2,O_3, CO	Gases in the air react to the surface of the sensor and alter its resistance.	
Optical particulate counter	PM	Detect particulate emissions by measuring the particulate lights scattered.	
Optical Sensor	CO, CO_2	Gases such as carbon monoxide and carbon dioxide are measured by measuring the absorption of infrared radiation.	
Photoionization detector	VOC	Volatile organic compounds and the resulting electric current is measured.	

III. COLLECTION AND PREPROCESSING OF DATA USING DATA FUSION

A. Collection of Data

Data Fusion (DF) methods have gained increasing attention from several societies over the past few years due to the growing demand to incorporate the huge amount of data generated by the systems. Data Fusion is widely interpreted according to the applications, innovations, and relevant populations. In a broader context, it refers to the gathering and analysis of various data from multiple sensors to get the properties of an ecosystem or entity. DF provides marked benefits over data from a single source. Besides the statistical benefit acquired by integrating data from the same source,

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

making use of these multiple sensors data will improve the precision where a volume can be measured and categorized. In this paper, a real-time air pollution data are collected from 3 different air quality monitoring sites. The data were acquired from the Central Pollution Control Board, Chennai region. Collected the meteorological data which also, will influence the concentrations of the air pollutants, including relative humidity(RH), air temperature (Temp), Wind Direction (WD), Wind speed(WS) and Pressure, etc. The locations of three air quality monitoring stations are Alandur Bus Depot, Chennai. Latitude 12.9099161 and Longitude 80.1076538. Manali, Chennai Latitude 13.164544 and Longitude 80.26285. Velachery res area Latitude 13.0052189 and Longitude 80.2398125. Sensor fusion plays a role here. During the data collection, different sensors are used to collect the data, for example, particle sensor detects the physical contaminant. Humidity and temperature are detected using thermal sensors. Gases like O2, NO2, SO2, CO2, CO are monitored using gas sensors.□

B. *Preprocessing*

After collecting the data, Heterogenous data fusion is performed which means the different types of data from different types of sensors. Fuse all three different location data into one. However, at some hours the original data may have duplicate records or missing incomplete values. Because of the huge amount of fluctuations and situations involved, which produce noise in timelines, the collected time series had to be revised and renewed. So as per the fusion guidelines, preprocessed data must be clustered or classified then filter the data and make a decision based on that. So here we proposed an LDA for clustering. After a classification Random forest tree is proposed for decision making.

Here we are analyzing the Pollution level for three different locations, hence here we do not require a particular date or time.

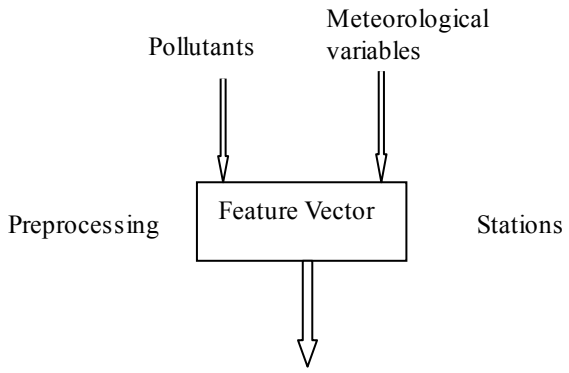


Fig. 2. Preprocessing for all the three stations

Proposed a system for air quality monitoring and tracking of the Chennai. Meteorological variables At the data fusion stage wind speed and wind direction are used as a deciding factor. Fig 3,4,5,6,7and 8 shows particle matters with different gases and meteorological variables. A data fusion framework provides many benefits over individual sensors, for instance, improved measurement through multiple measurements of a physical phenomenon when it comes to the simple task of using and transmitting data to a specific task.

IV. FEATURE SELECTION TECHNIQUE

A. *Linear Discriminant Analysis(LDA):*

LDA is a dimensionality reduction technique. It is used to reduce the dimensions. It is used to find differences in groups and partition two or more classes. LDA is one of the clustering techniques which is used to group the data and also for partitioning data and gives the best accuracy than compared to other algorithms. Apply LDA over the equivalent dataset twice with various jobs each time. In the primary methodology, LDA will fill in as a classifier and posterity it will decrease the dimensionality of the dataset and a neural system will play out the grouping task, the consequences of the two methodologies will be analyzed a short time later.

LDA makes a few rearrangements about the information:

1. That your information is Gaussian, that when plotted, every element is molded like a chime bend.
2. That there is a similar shift each time, that estimates of each element fluctuate across the mean by a similar amount of all items considered.

The LDA model gages the mean for each class with these presumptions, and adjustments from your knowledge. In the case of two groups in the univariate (single knowledge variable) it is anything but hard to understand this. In the traditional direction, the mean (m) estimate of each information (xi) for each class (ki) can be calculated, by dividing aggregate qualities by the absolute number of qualities.

$m = 1/nki * sum(xi)$

For class k, where m is the mean estimate of x, nki is the sum of occasions with class k. The adjustment is defined as the standard squared differentiation of each reward from the mean for all groups.

$sig^2 = 1 / (n-Ki) * sum((xi– m)^2)$

Where sig^2 is the variance of all inputs (xi), n is the number of instances, Ki is the number of groups, and m is the mean of xi.

B. *Predictions made with LDA:*

LDA does the prediction process by estimating the likelihood that it will be an input set belonging to each group. The group with the highest amount of probability is the class of output, and the process of prediction is made. The model here makes use of the Bayes Theorem to estimate probabilities. It is used to calculate the probabilities of the

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

given output class (ki), the input (xi) using each class's probability, and the likelihood of each class's data.

P(Y=x|iX=xi) = (Pok * fki(xi)) / sum(PIl * fl(xi))

Where Pok refers to the basic likelihood of each class (ki) found in the training. In Bayes' theorem, this is called the prior probability

Pok = nki/n

The above f(xi) is the approximate likelihood of xi belonging to a given class. For f(xi) a Gaussian distribution function is used. By solving the above equation, we get

D(xi) = xi * (m/sig^2) – (m^2/(2*sig^2)) + ln(Pok)

D(xi) is the discriminating function for class ki given with input xi, from which all them, sig2, and Pok are obtained.

Firstly, all the characteristics of that dataset that your work will be guiding you about the decision of applying LDA is a linear classifier or algorithm for the reduction of dimensionality to perform classification. The prior aim of, In fact, Linear Discriminant Analysis is an example of classes that slowly transfer them into a specific feature space, if the dataset is linearly separable, applying LDA as a classifier is the only way to achieve better results. Next, if the dataset is no longer linear separable, the LDA can seek as much as possible to arrange the data set in another space as the optimal linear separability, although it may still be an example of class duplication due to non-linear data characteristics. In this case, you have to need to use another classification technique to deal with nonlinear data, such as a neural network model with layers that are hidden, a neural network with radial basis functions that are available.

C. *Random Forest*

It is one of the managed arrangement calculation. It orders the huge dataset which contains enormous size. It is a choice calculation that is utilized for the group to give a precis ˆ cast. The irregular timberland calculation is a mana¡ arning model; it utilizes marked information to "realize" how to arrange unlabeled information. This is something contrary to the K-implies Cluster calculation. The Random Forest Algorithm can be used to resolve relapse issues as well as structural problems, making it an assorted model that engineers usually use. While performing the Random Forest algorithm, based on data classification, you must be informed that you are using the Gini index or the algorithm used to validate how the decision tree branch nodes are used.

$$Gini = 1 - \sum_{i=1}^{C} (p_i)^2$$

This method uses the class and probability to calculate Gini on a node for each branch, which branches are more likely to occur. Here pi is the class frequency you see in the dataset and c shows the number of classes. Entropy can also be used to determine how nodes in a decision tree can branch out.

$$Entropy = \sum_{i=1}^{C} -p_i * \log_2(p_i)$$

A single decision tree must be understood before we truly understand the random forest. We need to consider the variations between the present node, branch, and leaf and how different formulae can be used to finish the decision-making process. The random forest algorithm would be useful for all types of data sets, whether it's regression or classification data if performed correctly. It is simple to use, fast to train and gives an accurate image of the decision trees that it uses.□

V. RESULTS AND DISCUSSION

Real-time data is gathered from a different place and different time. We performed data analysis and gives the prediction which gives much accuracy. After analyzing the data from the datasets, the locations are identified as dangerous ones based on most severity and weather conditions. Fig 3-7 gives the various gas level based on PM2.5.

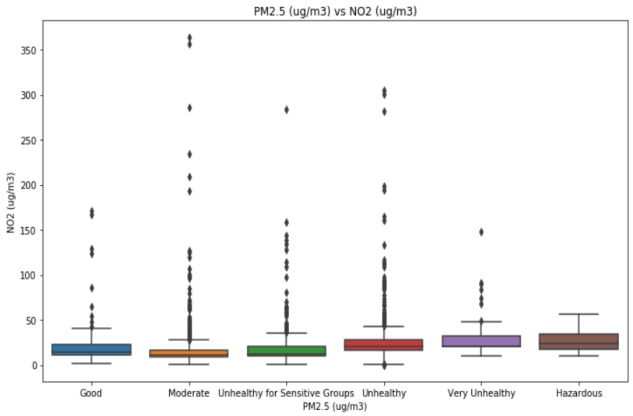


Fig. 3. PM2.5 vs NO2

In fig.3 PM2.5vs NO2 is shown. NO2 affects the lungs and enhances breathing infections which may lead to asthma problems.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

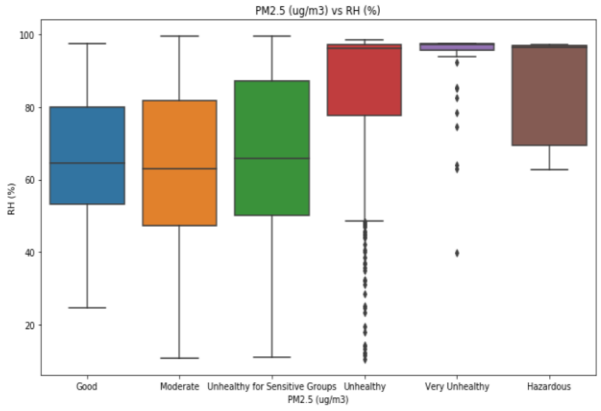


Fig. 4. PM2.5 vs RH

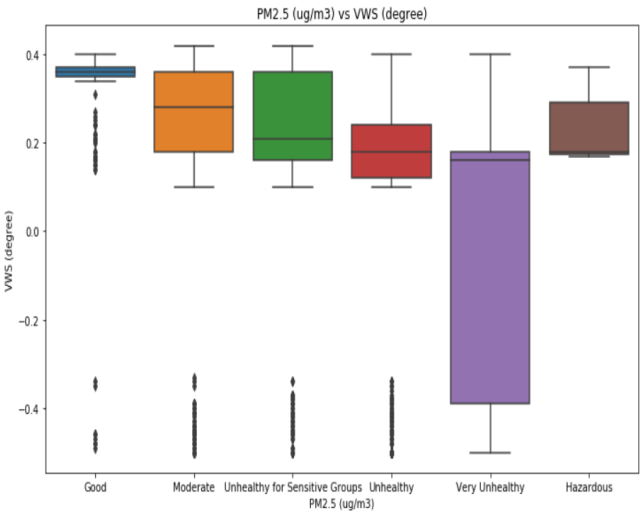


Fig. 5. PM2.5 vs VWS

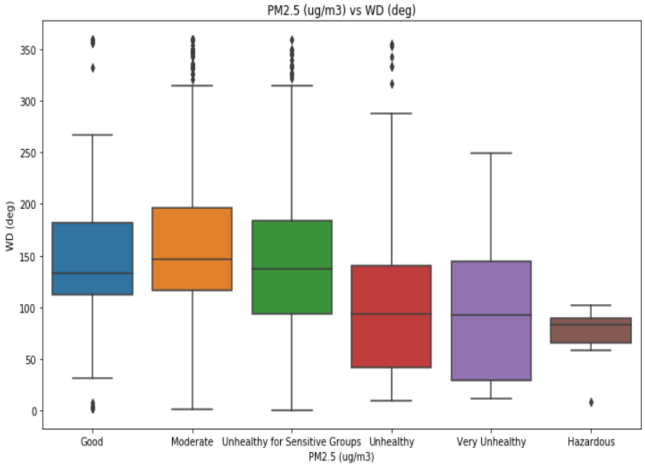


Fig. 6. PM2.5 vs WD

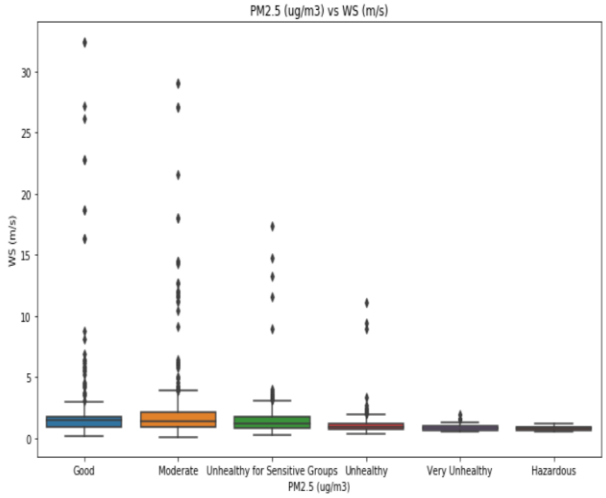


Fig. 7. PM2.5 vs WS

The direction of the wind is an important parameter that affects PM2.5. The wind carried an equal amount of contaminants from different directions. Fig5-7 gives the PM2.5 concentration level based on wind speed, wind direction, and relative humidity.

VI. CONCLUSION

The proposed work can predict the gas level by combining multiple sensor data from multiple cities and predict the pollution level for different gases using LDA and Random forest tree. We have collected data from the Central Pollution Control Board. Analyzed the data using python.

ACKNOWLEDGMENT

We are thankful to the Sathyabama Institute of Science and Technology management for providing all the requisite facilities to prepare this manuscript.

REFERENCES

[1] Yunliang Chen, Lizhe Wang , Fangyuan Li , Bo Du , Kim-Kwang Raymond Choo ,Houcine Hassan, Wenjian Qin, Air quality data clustering using EPLS method R, Information Fusion(2016)

[2] Eiman Kanjo a, Eman M.G. Younis b, Nasser Sherkat a, Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach, Information Fusion 40(2018) pp 18-31

[3] Wenxiu Ding a , Xuyang Jing a , Zheng Yan , Laurence T. Yang c , A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion, Information Fusion 51 (2019) 129–144

[4] Jia Liu, Tianrui Li, Peng Xie, Shengdong Du, Fei Teng , Xin Yang, Urban big data fusion based on deep learning: An overview, Information Fusion 53 (2020) 123–133

[5] Hamed Karimian, Qi Li, Chunlin Wu, Yanlin Qi, Yuqin Mo, Gong Chen, Xianfeng Zhang, Sonali Sachdeva, Evaluation of Different Machine Learning Approaches to Forecasting PM 2.5 Mass Concentrations, *Aerosol and Air Quality Research*, 19: 1400–1410, 2019

[6] R. O. Sinnott and Z. Guan, "Prediction of Air Pollution through Machine Learning Approaches on the Cloud," 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT), Zurich, 2018, pp. 51-60.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

[7] Air Pollution Detection Using Multisensor Data Fusion , Md Faruck , Punyasha Chatterjee, Workshops ICDCN '18: Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and NetworkingJanuary 2018 Article No.: 24 Pages 1–2

[8] Dixian Zhu, Changjie Cai, Tianbao Yang and Xun Zhou, A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization, Big Data Cogn. Comput. 2018, 2, 5; doi:10.3390/bdcc2010005

[9] A.Sivasangari, P.Ajitha, K.Indira, Air Pollution Monitoring and Prediction using Multi View Hybrid Model, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 6S4, April 2019

[10] Mrs. A. Gnana Soundari MTech, (PhD) , Mrs. J. Gnana Jeslin M.E, (PhD) , Akshaya A.C, INDIAN AIR QUALITY PREDICTION AND ANALYSIS USING MACHINE LEARNING, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 11, 2019

[11] Kostandina Veljanovska, Angel Dimoski , Air Quality Index Prediction Using Simple Machine Learning Algorithms International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Volume 7, Issue 1, January - February 2018

[12] Aditya C R, Chandana R Deshmukh , Nayana D K , Praveen Gandhi Vidyavastu, Detection and Prediction of Air Pollution using Machine Learning Models, International Journal of Engineering Trends and Technology (IJETT) – volume 59 Issue 4 – May 2018

[13] R.M.Gomathi, J.Martin Leo MAnickam, "Energy Efficient Shortest Path Routing Protocol for Underwater Acoustic Wireless Sensor Network ",Wireless Personal Communications, 98(1), pp. 843–856, Jan 2018.

[14] J.M. Barron-Adame, M.G. Cortina-Januchs and A. Vega-Corona," Data Fusion and Neural Network Combination Method for Air Pollution Level Monitoring", 978-1-4244-3760-3/09/$25.00 c 2009 IEEE, 2009 7th IEEE International Conference on Industrial Informatics (INDIN 2009)

[15] Andrew Fox, B. V. K. Vijaya Kumar, Jinzhu Chen, Fan Bai "Multi-Lane Pothole Detection from Crowdsourced Undersampled Vehicle Sensor Data" IEEE Transactions on Mobile Computing, Vol. 16, Issue. 12, 2017, 3417-3430.

[16] Jeong-Joo Kim, Su-il Choi, "Implementation of Pothole Detection System Using 2D LiDAR", International Conference on Electronics, Information, and Communication,2017, 607-610.

[17] https://ec.europa.eu/environment/air/pdf/Brochure%20lowercost%20sensors.pdf

[18] [2] Andrew Fox, B. V. K. Vijaya Kumar, Jinzhu Chen, Fan Bai "Multi-Lane Pothole Detection from Crowdsourced Undersampled Vehicle Sensor Data" IEEE Transactions on Mobile Computing, Vol. 16, Issue. 12, 2017, 3417-3430.

[19] Jeong-Joo Kim, Su-il Choi, "Implementation of Pothole Detection System Using 2D LiDAR", International Conference on Electronics, Information, and Communication, 2017, 607-610.

[20] Wang, Jianhua & Ogawa, Susumu. (2015). Effects of Meteorological Conditions on PM2.5 Concentrations in Nagasaki, Japan. International journal of environmental research and public health. 12. 9089-101. 10.3390/ijerph120809089.

Authorized licensed use limited to: Sathyabama Institute of Science and Technology. Downloaded on January 31,2023 at 14:59:59 UTC from IEEE Xplore. Restrictions apply.

An official website of the United States government
Here's how you know

FULL TEXT LINKS

J Nanosci Nanotechnol. 2020 May 1;20(5):2902-2910. doi: 10.1166/jnn.2020.17440.

Actinobacterial-Mediated Fabrication of Silver Nanoparticles and Their Broad Spectrum Antibacterial Activity Against Clinical Pathogens

Poongazhalselvan Devagi [1], Thirunavukarasu Chitrikha Suresh [2], Revathi Vedachalam Sandhiya [1], Mohan Sairandhry [1], Selvaraj Bharathi [1], Palanivel Velmurugan [1], Manikkam Radhakrishnan [3], Thangavelu Sathiamoorthi [4], Gopal Suresh [1]

Affiliations

PMID: 31635627 DOI: 10.1166/jnn.2020.17440

Abstract

The aim of this study is to fabricate silver nanoparticles (AgNPs) using actinobacterial strain isolated from lawn soil. Among six isolates, one isolate named AS-3 was potent in AgNPs production; hence it was identified deployed on gene sequence (16S rRNA) as *Streptomyces spongiicola* AS-3 (99.8% similarity). Actinobacteria mediated synthesized AgNPs were analyzed using UV-visible spectroscopy (UV-Vis), which showed a Surface Plasmon Resonance (SPR) at around λ = 443 nm. Scanning electron microscopy (SEM) analyses revealed the occurrence of predominant spherical AgNPs with polydispersed, with an average size of 22 nm. Energy-dispersive X-ray spectroscopy (EDS) established the existence of silver component. While the Fourier transforms infrared spectroscopy (FTIR) evidenced the occurrence of proteins as the bio reduction and topping agents over the AgNPs. X-ray diffraction (XRD) examination confirmed the obtained AgNPs were in crystalline planes of the face centric cubic. The *S. spongiicola* AgNPs antibacterial activity showed a broad spectrum antibacterial action against *Staphylococcus aureus, Bacillus cereus, Escherichia coli, Pseudomonas aeruginosa, Vibrio cholera, Shigella* sp., and *Salmonella typhi* were confirmed by disc diffusion test and MIC analysis.

Supplementary concepts

Streptomyces spongiicola

LinkOut - more resources

Full Text Sources

Ingenta plc

Research Materials

NCI CPTC Antibody Characterization Program

Received: 23 August 2020 | Revised: 17 September 2020 | Accepted: 20 September 2020

DOI: 10.1111/are.14942

Check for updates

REVIEW ARTICLE

Aquaculture Research WILEY

Actinobacteria and their bioactive molecules for anti-WSSV activity: A mini review

Ravindran Namitha[1] | Raguraman Vasantharaja[1] | Manickam Radhakrishnan[2] | Kasivelu Govindaraju[1]

[1]Centre for Ocean Research (DST-FIST Sponsored Centre), MoES - Earth Science & Technology Cell, Sathyabama Institute of Science and Technology, Chennai, India

[2]Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology, Chennai, India

Correspondence
Kasivelu Govindaraju, Centre for Ocean Research (DST-FIST Sponsored Centre), MoES - Earth Science & Technology Cell, Sathyabama Institute of Science and Technology, Chennai-600 119, India.
Email: govindtu@gmail.com

Funding information
Ministry of Earth Sciences (MoES), Grant/Award Number: (MoES/11-MRDFIESTC-MEB (SU)/2/2014 PCIII)

Abstract

Shrimp products are the best seafood commodity by value and confer about 17% of the total internationally traded fishery products. White Spot disease caused by white spot syndrome virus (WSSV) poses massive economic destruction globally in the shrimp aquaculture industry. Lack of proper treatment methods for WSSV infection puts shrimp farmers in trouble and increases the expenditure for the structure and maintenance of shrimp culture systems. The current review explores the significance of actinobacteria and their importance as economically and biotechnologically valuable prokaryotes that can produce a wide range of bioactive secondary metabolites and hence, develop lead compounds against White Spot disease.

KEYWORDS

Actinobacteria, aquaculture, secondary metabolites, WSSV

1 | INTRODUCTION

Aquaculture is the rapidly growing sector in the food and animal industry (Gephart et al., 2020). Globally, aquaculture and fisheries industries remain prime sources of food and nutrition, income generation and livelihoods for hundreds of millions of people (Subasinghe et al., 2009). As stated by FAO, aquaculture is referred as farming of aquatic organisms in coastal and inland areas involving interventions in the rearing conditions to improve production. According to the United Nations Food and Agriculture Organization, the global human population will consume 30 million tons of fish by 2030 (Gephart et al., 2020). The global per capita seafood supply reached a new record high of 20 kg in 2014 (FAO, 2016). The performance over the last 30 years, and sufficiently stable catches from capture fisheries, clearly indicates that the fisheries sector's future growth will come mainly from aquaculture (FAO, 2016). Aquaculture is one of the best possible sustainable solutions for food security and increased dietary nutrition in developing regions (Sinead, 2013). Production in aquaculture sector comes primarily from fish ponds (freshwater and marine water), fish cages, and fish pens (freshwater and marine water), and mariculture (FAO, 2001). According to the FAO report, 59.6 million people seem to be involved in the primary sector of capture fisheries and aquaculture in 2016, with 19.3 million people engaged in aquaculture and 40.3 million people active in fisheries (FAO, 2018).

Globally, over one billion population depends on fish as their major source of animal protein. In many aquaculture predominant areas, the most remarkable species of fish produced are utilized for local consumption. The technological advancement in aquaculture has tremendously increased (more than doubled) the animal protein consumption per capita in developing countries in the last 30 years (Sinead, 2013). In aquaculture, serves as sufficient nutritional food for those living in lower socio-economic households to ensure their health and well-being. In developing states, the fish farming practices have increased exponentially, with more agricultural farmers discovering that aquaculture's integration into their production systems leads to improved land and yield productivity (Sinead, 2013). Meanwhile, shrimp farming has become a significant aquaculture activity and attractor of investment over the past two to three decades. The global market for shrimps estimated to USD 45 billion in 2018, and it is projected to register a CAGR of 5.2% during the forecast period, 2019–2024. The world's leading shrimp producers are china and Thailand, accounting for nearly 75% of the global shrimp production. Shrimp production in Asia has improved since 2015,

 © 2020 John Wiley & Sons Ltd

reaching 3.75 million MT in 2018 and potentially 4.00 million MT in 2021 (Figure 1) (Anderson et al., 2019).

2 | WHITE SPOT SYNDROME VIRUS (WSSV)—AN OVERVIEW

The fastest aquaculture growth is expected for shrimp in the coming years, with the most massive expansion in India. In parallel, the popularity of shrimp farming is increasing globally. There are currently 50 countries worldwide, focusing on two major regions, Asia and the America (Jory, 2018). Apart from lack of land and regulatory guidance, India's shrimp farming industry had a rapid growth during the 1990s (Salunke et al., 2020). In 1994, a disease affecting the most widely used shrimp species, *Penaeus monodon*, became widespread and negatively impacted India's entire industry (Salunke et al., 2020). White Spot Syndrome disease is a highly virulent and contagious disease of crustaceans, especially affecting shrimps caused by WSSV (Dashtiannasab, 2020). The double-stranded DNA virus belongs to the family *Nimaviridae*, genus *Whispovirus*, which causes white spot syndrome in a wide range of aquatic crustaceans (Wang et al., 2019). The WSSV infection of penaeid shrimp is highly lethal causing high mortality (100%) in shrimp, since its emergence in 1990s. Outbursts of this disease have resulted in the massive wipeout of shrimp populations all over the world (Dashtiannasab, 2020).

White spot syndrome virus is one of the most devastating pathogens, causing considerable losses to the shrimp farming sector worldwide, it can quickly spread in farming sites (Dashtiannasab, 2020). It shows both vertical and horizontal modes of transmission in the natural environment. Lack of effective treatment methods against this disease adds to the concern. The occurrence of white spots (ranging from 0.5 to 3.0 mm in diameter) on the exoskeleton, appendages and inside the epidermis on shrimp are the major physiological symptoms of WSSV infection (Sanchez-Paz, 2010). Reports suggest that clinical signs' appearance follows the high and rapid mortality of the infected population in a short period. Virions of WSSV are ovoid/ellipsoid to bacilliform in shape, have a regular symmetry, and measure 80–120 nm in diameter and 250–380 nm in length with a flagella-like-extension at one end of the virion. Studies indicate that the WSSV virion consists of at least 58 structural proteins (Sánchez-Paz, 2010). The viral proteins VP28 and VP26 are the major envelope proteins contributing to 60% of the envelope among the other 35 different proteins present in the viral envelope. Among the other viral proteins, VP28 and VP26 form trimers in the viral envelope, providing insight into the role of viral proteins in causing infection in host cells (Figure 2) (Sanchez-Paz, 2010). The viral infection in host cells was mediated through these structural viral proteins via recognition and attachment to receptors in host cell surface. However, detailed molecular mechanism of WSSV-host interactions will provide efficient treatment strategies against WSSV (Li et al., 2019). In future, studies should focus on determining the actual mechanism of viral infection and pathogenesis on the populations of crustaceans. Moreover, several therapeutic measures are available such as seaweed-based sulphated polysaccharides for the effective management of WSSV in shrimps (Vasantharaja et al., 2020). In addition, vaccination strategies including DNA vaccines were developed to mitigate WSSV infection in shrimps. In spite of development, the safety issues are the major concern in developing vaccines than its therapeutic potential (Haq et al., 2012). Thus, there is an immediate need for alternative treatment methods to prevent and control WSSV infection.

3 | ACTINOBACTERIA AND THEIR IMPORTANCE

Phylum actinobacteria remain among the largest in the bacterial domain with the most enormous industrial and economic significance. Actinobacteria include the largest group of Gram-positive bacteria with high guanine plus cytosine content in their genomic DNA. Members of this phylum are existing in diverse morphological forms ranged from simple unicellular cocci, rods to complex filamentous structures (Figure 3). According to the recent report by (Salam et al., 2020) based on the refined phylogeny of the 16S rRNA sequences, all the members of the 425 genera of the phylum 'Actinobacteria' are rearranged into six classes, 46 orders and 79 families, including 16 new orders and ten new families. The six actinobacterial classes include the Acidomicrobiia, Actinomycetia, Coriobacteria, Nitriliruptoria, Rubrobacteria and Thermoleophilia. In general, most actinobacterial genera are aerobic, whereas some

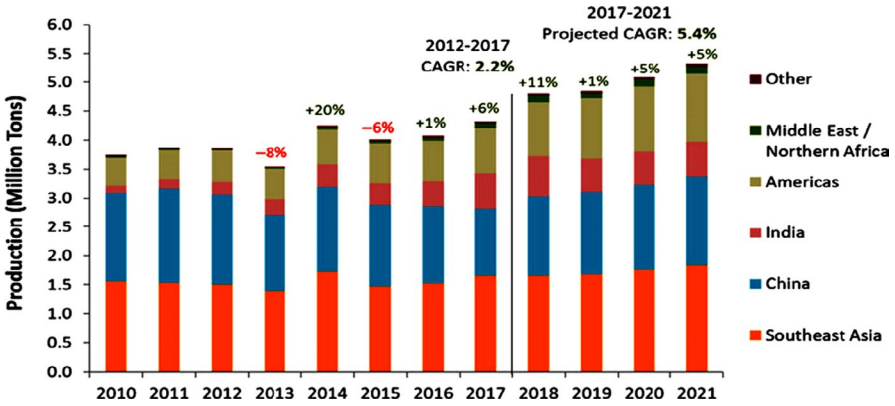


FIGURE 1 Shrimp aquaculture production by region (Source: Anderson et al., 2019)

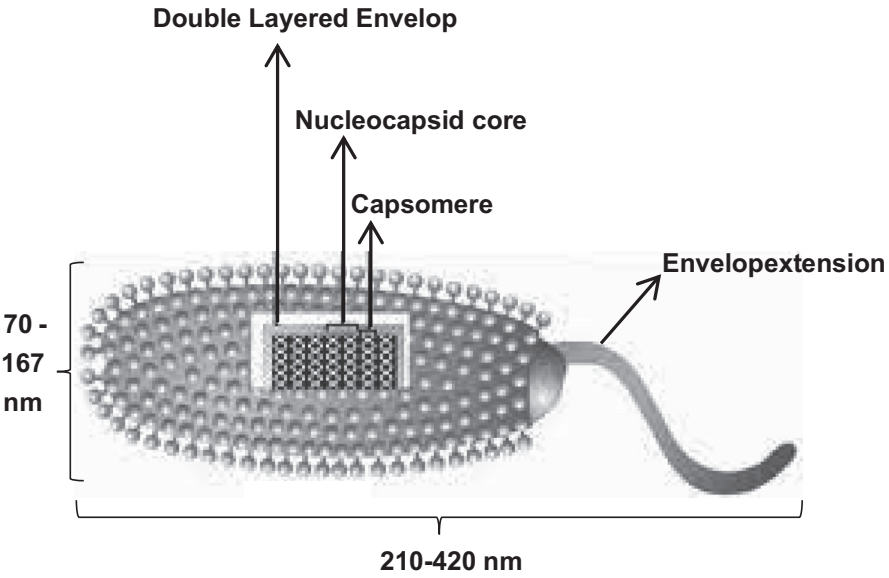


FIGURE 2 Schematic representation of the structural morphology of the white spot syndrome virus (WSSV) particle (Source: Sanchez-Paz, 2010)

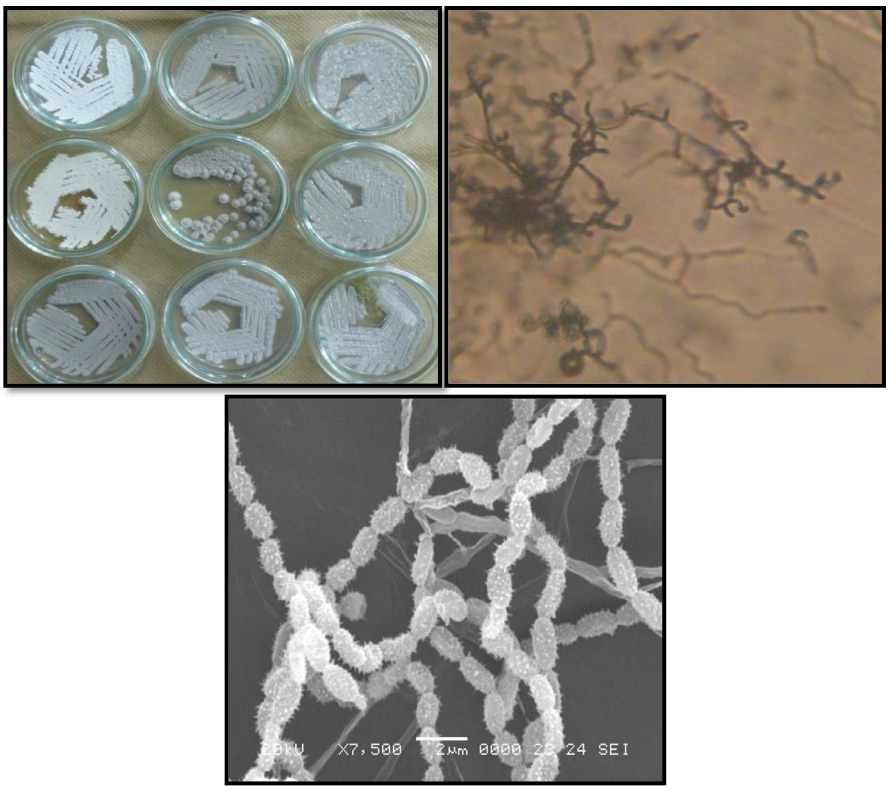


FIGURE 3 Cultural and microscopic appearance of Streptomyces/ actinobacteria

genera tend to be facultative or obligatory anaerobes. They are metabolically chemoheterotrophic, and they utilize diverse substrates, including complex polymers for energy. Most actinobacteria are free-living saprophytes in numerous habitats such as typical terrestrial and aquatic environments and living hosts like plants, animals and insects as symbionts. However, the distribution of several actinobacterial genera in the extreme habitats like the greatest depths of the ocean, Antarctica, Caves, high altitude Mountain and desert soil was also reported (Barka et al., 2016; Berdy, 2012). Also, some are pathogens for humans, animals and plants.

Among the several actinobacterial genera, the genus *Streptomyces* is dominant in any ecosystem. Other than the genus *Streptomyces*, all other genera that come under the phylum actinobacteria are called rare actinobacteria. As per the early literature, unique actinobacterial genera are difficult to isolate by routine isolation procedures and less frequently distributed. Actinobacteria are proficient in sustaining into different forms of ecological niches due to their physiological, biosynthetic and enzymatic potential (Salaria et al., 2017). According to recent literature, rare species of actinobacteria are relatively abundant in various habitats, especially in extreme ecosystems like the marine desert and caves. Actinobacteria can be retrieved in huge numbers by using

diverse suitable isolation protocols (Amin et al., 2020). Actinobacteria can also be isolated from marine sediments, seawater, aquatic plants and animals. Marine actinobacteria, being the most abundant source of novel and essential compounds, represent a genus of less exploited marine microbes (Rasool & Hemalatha, 2017).

Actinobacteria, a largest phylum under the bacterial domain, has received remarkable attention mainly due to its ability to produce an immense array of bioactive secondary metabolites with complex chemical structures and different biological properties (Barka et al., 2016). In the field of medicine, more than 65% of antibiotics was developed based on actinobacteria, out of which over 10,000 bioactive compounds are produced by the members of the genus *Streptomyces* (Berdy, 2012; Lee et al., 2018; Valliappan et al., 2014). The genus *Streptomyces* remains in the focus of bio-prospecting, essentially due to its unrivaled ability in synthesizing a huge array of bioactive compounds with different bioactivities (Lee et al., 2018; Ser et al., 2017; Tan et al., 2016). Among the different microorganisms, actinobacteria are undoubtedly the largest producers of secondary metabolites (Berdy, 2005). Actinobacterial secondary metabolites exhibit diverse biological activities such as antibacterial, antifungal, antiviral, anti-cancer, anti-diabetic and immunosuppressive (Radhakrishnan et al., 2019).

The antibiotics synthesized by actinobacteria are classified into different types based upon their structure, such as aminoglycosides, ansamycins, anthracyclines, β-lactam, macrolides and tetracycline (Amin et al., 2020). However, the mechanism of action seems to be different for individual metabolites produced by rare actinobacteria, such as inhibition of cell wall synthesis, cell membrane damage, inhibition of DNA and RNA synthesis, protein synthesis inhibition and inhibition of essential enzymes required for metabolism. Interestingly, in recent years, the search for extreme ecosystems results in the isolation of promising actinobacterial strains that produce novel bioactive metabolites.

Actinobacteria are extensively involved in the production of industrial enzymes like amylases, cellulases, chitinases, proteases, xylanases and others (Prakash et al., 2013). The enzymes derived from actinobacteria seem to have substantial commercial potential, mainly due to their nature of stability and unusual substrate specificity (Salaria et al., 2017). Actinobacteria also play a significant role in the production of other industrial important molecules like bio-surfactants and bio-emulsifiers. Different types of bio-surfactants, like glycolipids, pentasaccharide lipids, trehalosetetrae esters, lipopeptides and peptidolipids are produced by different actinobacterial genera such as *Streptomyces*, *Nocardia* and *Rhodococcus* sp. Actinobacteria are capable of performing a diverse microbial conversions of organic compounds, where it aided as necessary factor in the conversion of biodegradation of pollutants in soil and water, oil spills, and chemical waste decomposition. Moreover, members of the genera *Nocardia* and *Streptomyces* play a vital role in the chemical modifications of complex compounds of natural and synthetic origin (Salaria et al., 2017). These strains also can degrade certain pesticides and herbicides in soil. Actinobacteria enhances soil fertility by involving key processes including nutrient cycling, decomposition, development of beneficial soil humus, and assisted in control of plant pathogens, insects and weeds (Suarez-Moreno et al., 2019). Free-living actinobacteria produce plant-growth-promoting hormones like auxins and gibberellin-like compounds (Bloemberg & Lugtenberg, 2001; Persello-Cartieaux et al., 2003). They also play a substantial role in bioremediation by eliminating the toxic xenobiotics in the environment and contributing to the recycling of organic carbon. Marine actinobacteria have been reported to be the potential sources for the production of enzyme inhibitors, including glucosidase, N-acetyl-D-glucosaminidase, pyroglutamyl peptidase and amylase inhibitors (Manivasagan et al., 2015).

4 | ACTINOBACTERIA VERSUS ANTIVIRAL DRUG DISCOVERY

In the early 1950s, there are no active antiviral agents, mainly due to the lack of screening methods for screening antiviral substances. Screening of antiviral agents has been done by animal experiments using mice or by adopting the chick embryo method. However, these methods are not quantitative and not so efficient at that time to screen a large number of materials in short time duration. Therefore, tissue culture has been proposed as a simple in vitro method for testing many extracts and compounds for antiviral properties (Raveh et al., 2013).

Due to the advent of fermentation and extract preparation and virus cultivation and preservation techniques, there are several actinobacterial extracts and compounds screened using the in vitro cell culture platform. However, still, the number of antiviral compounds reported from actinobacteria are very less when compared to the number of antibacterial and antifungal compounds reported from them. In recent years actinobacteria from understudied/extreme ecosystems or sources are found to produce antiviral compounds (Kim et al., 2016). The antiviral activities of pentalactones and pyrrole-2-carboxylic acid had been assayed and described against herpes simplex virus (McLaren et al., 1983). Fattiviracin, an antiviral agent from *Streptomyces microflavus*, has also been reported to inhibit herpes simplex virus 1 (HSV-1) and human immunodeficiency virus (HIV) type 1 (Habib et al., 2001). Previously, (Nakagawa et al., 1985) reported that pentalactones isolated from Streptomyces M-2718 fermented broth has been found to have potential activity against DNA viruses. Another study showed guanine-7-N-oxide isolated from *Streptomyces* sp. has the ability to inhibit fish herpes virus (Onchorhynchus Masou Virus), rhabdovirus (Infectious Hematopoietic Virus) and a fish virus (Infectious Pancreatic necrosis Virus) (Hasobe & Saneyoshi, 1985). The current review focuses on the present status, possibilities and challenges of exploring actinobacteria and their metabolites to manage WSSV infection in aquaculture.

5 | EFFECTS OF ACTINOBACTERIAL METABOLITES AGAINST WSSV

Several decades ago, actinobacterial research focused on hunting bioactive metabolites, a wide variety of compounds with various structural and functional properties are identified and characterized. As per the study conducted in 2017, India represents 17.24%

of the WSSV publications globally, and remain second among the 15 countries in terms of the number of publications in this period (Bhoomaiah et al., 2020). This article provides insights into the bioprospecting of actinobacteria and the novel compounds developed against the white spot disease.

Recently, a study claim that the benzoic derivatives characterized by the haloalkaliphilic actinomycete *Streptomyces castaneoglobisporus* AJ9 may be interacted with the outer and nucleocapsid proteins of the WSSV and can affect the transcription leading to WSSV multiplication (Adlin Jenifer et al., 2018). Another study identified the antiviral property of six actinobacterial isolates from the coastal waters of Cochin against WSSV in Penaeid shrimps (Kumar et al., 2006). Extracts of 25 strains of actinobacterial species were incorporated into feed additives and tested for their ability to reduce infection among cultured shrimps. The crude extract integrated pellet feed of 10 g was fed to the post-larvae of *Penaeus monodon* for 2 weeks and then challenged with WSSV. Six feeds among them, S.A. 2, S.A. 8, SL 27, SL 33, SL 39 and SL 85, were observed to be the most potent candidates with the highest post-challenge survival of above 50% and three of them, S.A. 2, S.A. 8 and SL 85, showed post-challenge survival of above 70%.

Studies also reported the potential of marine actinobacteria in improving disease resistance against White Spot viral disease in *Penaeus monodon*. Actinobacterial strains (77 no.) isolated from sediment samples collected from the West coast and the Southwest coasts of India have been extensively studied on anti-WSSV activity (Nair, 2012). This study reported that a medicated diet (20 g) for the animal was prepared by incorporating concentrated actinomycete broth (5 ml of each strain). After the secondary screening of the actinobacteria for the antiviral property, seven strains depicting maximum post-challenge survival (more than 40%) were selected. The strains with maximum inhibition percentage was observed in L45 *Streptomyces griseus* (60.27%), L56 *Streptomyces* sp. (54.6%), L10 *Streptomyces* sp. (51.1%), B451 *Streptomyces fradiae* (49%), L33 *Streptomyces flavidofuscus* (47.2%) L35 *Nocardia nova* (40.2%) and L27 *Brevibacterium linens* (40%). The biomass and concentrated supernatants of the seven actinobacterial strains were separately combined into a commercial diet. The supernatant incorporated diets were showed better survival rate of *P. monodon* species when challenged with WSSV. Among the groups, L45, L56 and B451 showed maximum post-challenge survival rates. The antiviral properties of the methanol extract of the isolates were found to be lesser effective when compared to the whole broth and also these extracts (L27, L35, L45, L56 and B451) showed 40% survival. Another study reported, 64 actinobacterial isolates collected by from different environments, a specific strain designated as CAHSH–2 exhibited potent anti-WSSV activity at low concentration (0.2 mg) per shrimp (Rajkumar et al., 2018). Antiviral activity of the extracts was tested in marine shrimp, *Litopeneaus vannamei* infected with WSSV was confirmed by RT-PCR, Western blot analysis and ELISA. Furthermore, shrimp injected with CAHSH – 2 treated WSSV showed a better survival rate (88.9 to 100%) than the shrimp treated with the extract of CAHSH – 2 (81.4 to 98.3%). CAHSH–2 was isolated from *Ocium tenuiflorum* and showed 99% similarity with *Streptomyces ghanaensis* and hence designated as *Streptomyces ghanaensis*—CAHSH–2. Further, extract was purified by silica gel column chromatography yielded two sub-fractions, F1A and F1B, of which F1B was found to be more productive and better survival rate. The active component was further characterized as Di-N-Octyl Phthalate based on GC–MS, FT-IR, and NMR analysis. Molecular docking studies revealed the maximum binding stability with VP28 and VP26, which increased the possibility of the compound to be a promising candidate against WSSV.

Manimaran et al., 2018 reported the anti-WSSV activity of 9(10H)-Acridanone compound derived from marine *Streptomyces fradiae* strain VITMK2 (*S. fradiae*) by screening 31 actinobacteria isolates collected from mangrove forest, Pichavaram, Tamil Nadu. Morphological, biochemical and molecular characterization of the strain led to the species identification *S. fradiae* strain designated as VITMK2. WSSV infected *Litopenaeus vannamei* treated with ethyl acetate extract of *S. fradiae* strain VITMK2 (500 mg/animal) showed a survival rate of 83.3%. The extract further purified by silica gel column chromatography yielded two fractions namely, C1 and C2. The WSSV infected shrimp treated with C1 fraction with different concentration (500 mg, 250 mg and 125 mg) showed survival rates of 88.89%, 83.33% and 55.56%, respectively, and also showed better antiviral activity. Molecular docking studies revealed least binding energy scores of −5.7 kcal/mol and −5.2 kcal/mol for VP26 and VP28 respectively. Docking study showed that the compound has strong interaction with the targets, and it can be effectively utilized as an active anti-WSSV agent for the control of WSSV infection. The available information points to the minimal attention paid to correlate the existing natural products research on actinobacteria with the evolving problems caused by white spot syndrome virus in aquaculture. It is the need of the hour to exploit the immense potential of these organisms and endeavour to abate the threats of the aquaculture industry.

6 | FUTURE PROSPECTS

- The recent advancements in molecular biology and omics approach will provide possible information on the mechanism and pathogenesis of WSSV in host cells.
- Intensive research will expand our knowledge about the fundamental nature of WSSV, its exact life cycle and its mode of infection to develop therapeutic measures.
- The substantial power of actinobacteria as the critical sources of bioactive secondary metabolites can quickly account for the discovery and description of salient compounds of commercial importance.
- Isolation and characterization of more varieties of the actinobacterial genera can lead to an ocean of possibilities to natural product research, and their further applications to control the WSSV.

7 | CONCLUSION

To explore the potential of actinobacteria, there is comparatively a lack in the novel antiviral compounds that are produced from actinobacterias. There is a need for new dimensional research in

actinobacteria to solve the problems prevailing in the aquaculture industry. Possible application of bioactive compounds from actinobacteria as immunostimulants by injections, oral administrations and administration in culture systems, increases the survival rates of Penaeid shrimps against WSSV. Elucidation of the mechanism of WSSV infection and the mode of action of the virus in the host helps develop effective therapeutics.

ACKNOWLEDGMENTS

The authors thank to Ministry of Earth Sciences (MoES)-Earth Science and Technology Cell (ESTC), Government of India, for financial assistance (MoES/11-MRDFIESTC-MEB (SU)/2/2014 PCIII). Authors thank management of Sathyabama Institute of Science and Technology (SIST) for its strong support to carry out research activities.

CONFLICT OF INTEREST

All the authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

GK planned and designed the organization of the manuscript. NR wrote the initial draft of the manuscript. VR reviewed the literature and added additional sections. RM wrote the part of the manuscript. VR and GK revised and edited the final manuscript. All authors read and approved the manuscript.

DATA AVAILABILITY STATEMENT

Data are available on request from the authors.

ORCID

Kasivelu Govindaraju https://orcid.org/0000-0002-9687-2174

REFERENCES

Adlin Jenifer, J., Remya, R., Velmurugan, S., Michaelbabu, M., & Citarasu, T. (2018). *Streptomyces castaneoglobisporus* AJ9, A haloalkaliphilic actinomycetes isolated from solar salt works in Southern India and its pharmacological properties. *Indian Journal of Geo Marine Sciences*, *47*(02), 475–488.

Amin, D. H., Abdallah, N. A., Abolmaaty, A., Tolba, S., & Wellington, E. M. H. (2020). Microbiological and molecular insights on rare *Actinobacteria* harboring bioactive prospective. *Bulletin of the National Research Centre*, *44*, 1–12. https://doi.org/10.1186/s42269-019-0266-8

Anderson, J. L., Valderrama, D., & Jory, D. E. (2019). GOAL 2019: Global shrimp production review. *Global Aquaculture Advocate*, 1–5.

Barka, E. A., Vatsa, P., Sanchez, L., Gaveau-Vaillant, N., Jacquard, C., Klenk, H. P., Clement, C., Ouhdouch, Y., & van Wezel, G. P. (2016). Taxonomy, physiology, and natural products of *Actinobacteria*. *Microbiology and Molecular Biology Reviews*, *80*(1), 1–43. https://doi.org/10.1128/MMBR.00019-15

Berdy, J. (2005). Bioactive microbial metabolites. *Journal of Antibiotics*, *58*(1), 1–26. https://doi.org/10.1038/ja.2005.1

Berdy, J. (2012). Thoughts and facts about antibiotics: Where we are now and where we are heading. *Journal of Antibiotics*, *65*, 385–395. https://doi.org/10.1038/ja.2012.27

Bhoomaiah, D., Krishnan, P., Kantharajan, G., Sangeeta, B., & Rajendran, K. V. (2020). A scientometric assessment of research on white spot syndrome virus (WSSV) in India vis-a-vis the world (1998–2017). *Aquaculture*, *520*, 734672. https://doi.org/10.1016/j.aquaculture.2019.734672

Bloemberg, G. V., & Lugtenberg, B. J. J. (2001). Molecular basis of plant growth promotion and biocontrol by rhizobacteria. *Current Opinion in Plant Biology*, *4*(4), 343–350. https://doi.org/10.1016/S1369-5266(00)00183-7

Dashtiannasab, A. (2020). White spot syndrome virus. *Emerging and Reemerging Viral Pathogens*, 717–728.

FAO (2001). *Production, accessibility, marketing and consumption patterns of freshwater aquaculture products in Asia: A cross-country comparison*. Fisheries Circular No. 973 FIRI/C973.

FAO (2016). *The State of World Fisheries and Aquaculture*. FAO.

FAO (2018). *The state of world fisheries and Aquaculture 2018-Meeting the sustainable development goals*. FAO, Licence: CC BY-NC-SA 3.0 IGO.

Gephart, J. A., Golden, C. D., Asche, F., Belton, B., Brugere, C., Froehlich, H. E., Fry, J. P., Halpern, B. S., Hicks, C. C., Jones, R. C., Klinger, D. H., Little, D. C., McCauley, D. J., Thilsted, S. H., Troell, M., & Allison, E. H. (2020). Scenarios for global aquaculture and its role in human nutrition. *Reviews in Fisheries Science & Aquaculture*, 1–17. https://doi.org/10.1080/23308249.2020.1782342

Habib, E. S. E., Yokomizo, K., Nagao, K., Harada, S., & Uyeda, M. (2001). Antiviral activity of Fattiviracin FV-8 against Human Immunodeficiency Virus Type 1 (HIV-1). *Bioscience, Biotechnology, and Biochemistry*, *65*(3), 683–685. https://doi.org/10.1271/bbb.65.683

Haq, B. A. B., Vignesh, R., & Srinivasan, M. (2012). Deep insight into white spot syndrome virus vaccines: A review. *Asian Pacific Journal of Tropical Disease*, *2*(1), 73–77. https://doi.org/10.1016/S2222-1808(12)60018-8

Hasobe, M., & Saneyoshi, M. (1985). On the approach to the viral chemotherapy against infectious hematopoietic necrosis virus (IHNV) *in vitro* and *in vivo* on salmonid fishes. *Fish Pathology*, *20*(2/3), 343–351. https://doi.org/10.3147/jsfp.20.343

Jory, D. E. (2018). Current production, challenges, and the future of shrimp farming. *Global Aquaculture Advocate*, 1–8.

Kim, S. H., Ha, T. K. Q., Oh, W. K., Shin, J., & Oh, D. C. (2016). Antiviral Indolosesquiterpenoid Xiamycins C-E from a halophilic Actinomycete. *Journal of Natural Products*, *79*(1), 51–58. https://doi.org/10.1021/acs.jnatprod.5b00634

Kumar, S. S., Philip, R., & Achuthankutty, C. T. (2006). Antiviral property of marine actinomycetes against white spot syndrome virus in penaeid shrimps. *Current Science*, *91*(6), 807–811.

Lee, L. H., Chan, K. G., Stach, J., Wellington, E. M. H., & Goh, B. H. (2018). Editorial: The search for biological active agent(s) from Actinobacteria. *Frontiers in Microbiology*, *9*, 824. https://doi.org/10.3389/fmicb.2018.00824

Li, C., Weng, S., & He, J. (2019). WSSV–host interaction: Host response and immune evasion. *Fish & Shellfish Immunology*, *84*, 558–571. https://doi.org/10.1016/j.fsi.2018.10.043

Manimaran, M., Rajkumar, T., Vimal, S., Taju, G., Abdul Majeed, S., Sahul Hameed, A. S., & Kannabiran, K. (2018). Antiviral activity of 9(10H)-Acridanone extracted from marine *Streptomyces fradiae* strain VITMK2 in *Litopenaeus vannamei* infected with white spot syndrome virus. *Aquaculture*, *488*, 66–73. https://doi.org/10.1016/j.aquaculture.2018.01.032

Manivasagan, P., Venkatesan, J., Sivakumar, K., & Kim, S. K. (2015). Actinobacterial enzyme inhibitors – A review. *Critical Review in Microbiology*, *41*(2), 261–272. https://doi.org/10.3109/1040841X.2013.837425

McLaren, C., Ellis, M. N., & Hunter, G. A. (1983). A colorimetric assay for the measurement of the sensitivity of herpes simplex viruses to antiviral agents. *Antiviral Research*, *3*(4), 223–234. https://doi.org/10.1016/0166-3542(83)90001-3

Nair, L. G. (2012). *Marine Actinomycetes as Source of Antiviral Agents and as Probionts for Penaeus monodon Culture Systems [Doctoral dissertation,*

Cochin University of Science and Technology]. Retrieved from http://hdl.handle.net/10603/4875.

Nakagawa, A., Tomoda, H., Hao, M. V., Okano, K., Iwai, Y., & Omura, S. (1985). Antiviral activities of pentalenolactones. *The Journal of Antibiotics*, *38*(8), 1114–1115. https://doi.org/10.7164/antibiotics.38.1114

Persello-Cartieaux, F., Nussaume, L., & Robaglia, C. (2003). Tales from the underground: Molecular plant-rhizobacteria interactions. *Plant, Cell and Environment*, *26*, 189–199. https://doi.org/10.1046/j.1365-3040.2003.00956.x

Prakash, D., Nawani, N., Prakash, M., Bodas, M., Mandal, A., Khetmalas, M., & Kapadnis, B. (2013). Actinomycetes: A repertory of green catalysts with a potential revenue resource. *BioMed Research International*, *2013*, 264020. https://doi.org/10.1155/2013/264020

Radhakrishnan, R., Premalatha, P., Shanmugasundaram, T., Gopikrishnan, V., Jerrine, J., Balagurunathan, R., & Dastager, S. G. (2019). Distribution and bioprospecting potential of Actinobacteria from Indian mangrove ecosystems. *Microbial Diversity in Ecosystem Sustainability and Biotechnological Applications*, 319–353.

Rajkumar, T., Manimaran, M., Taju, G., Vimal, S., Abdul, M. S., Kannabiran, K., Sivakumar, S., Kumar, K. M., Madhan, S., & Sahul Hameed, A. S. (2018). Antiviral viral compound from Streptomyces ghanaensis like strain against white spot syndrome virus (WSSV) of shrimp. *BioRxiv*, 340265.

Rasool, U., & Hemalatha, S. (2017). Diversified role and applications of marine actinomycetes in the field of biology. *World Scientific News*, *74*, 1–14.

Raveh, A., Delekta, P. C., Dobry, C. J., Peng, W., Schultz, P. J., Blakely, P. K., Tai, A. W., Matainaho, T., Irani, D. N., Sherman, D. H., & Miller, D. J. (2013). Discovery of potent broad spectrum antivirals derived from marine Actinobacteria. *PLoS ONE*, *8*(12), e82318. https://doi.org/10.1371/journal.pone.0082318

Salam, N., Jiao, J. Y., Zhang, X. T., & Li, W. J. (2020). Update on the classification of higher ranks in the phylum *Actinobacteria*. *International Journal of Systematic and Evolutionary Microbiology*, *70*(2), 1331–1355. https://doi.org/10.1099/ijsem.0.003920

Salaria, N., Sharma, S., & Sharma, S. (2017). *Actinomycetes: Potential and Applications*, *IV(VIII)*, 32–43.

Salunke, M., Kalyankar, A., Khedkar, C. D., Shingare, M., & Khedkar, G. D. (2020). A review on shrimp aquaculture in India: Historical perspective, constraints, status and future implications for impacts on aquatic ecosystem and biodiversity. *Reviews in Fisheries Science & Aquaculture*, *28*(3), 283–302. https://doi.org/10.1080/23308249.2020.1723058

Sanchez-Paz, A. (2010). White spot syndrome virus: An overview on an emergent concern. *Veterinary Research*, *41*, 43. https://doi.org/10.1051/vetres/2010015

Ser, H. L., Tan, L. T. H., Law, J. W. F., Chan, K. G., Duangjai, A., Saokaew, S., Pusparajah, P., Ab Mutalib, N. S., Khan, T. M., Goh, B. H., & Lee, L. H. (2017). Focused review: Cytotoxic and antioxidant potentials of mangrove-derived streptomyces. *Frontiers in Microbiology*, *8*, 2065. https://doi.org/10.3389/fmicb.2017.02065

Sinead, L. (2013). Fish for the future: Aquaculture and food security. *Future Directions International*, 1–8.

Suarez-Moreno, Z. R., Vinchira-Villarraga, D. M., Vergara-Morales, D. I., Castellanos, L., Ramos, F. A., Guarnaccia, C., Degrassi, G., Venturi, V., & Moreno-Sarmiento, N. (2019). Plant-growth promotion and bio-control properties of three Streptomyces spp. isolates to control bacterial rice pathogens. *Frontiers in Microbiology*, *10*, 290.

Subasinghe, R., Soto, D., & Jia, J. (2009). Global aquaculture and its role in sustainable development. *Reviews in Aquaculture*, *1*(1), 2–9. https://doi.org/10.1111/j.1753-5131.2008.01002.x

Tan, L. T. H., Chan, K. G., Lee, L. H., & Goh, B. H. (2016). *Streptomyces* bacteria as potential probiotics in Aquaculture. *Frontiers in Microbiology*, *7*, 79. https://doi.org/10.3389/fmicb.2016.00079

Valliappan, K., Sun, W., & Li, Z. (2014). Marine actinobacteria associated with marine organisms and their potentials in producing pharmaceutical natural products. *Applied Microbiology and Biotechnology*, *98*, 7365–7377. https://doi.org/10.1007/s00253-014-5954-6

Vasantharaja, R., Namitha, R., Koushalya, S., & Govindaraju, K. (2020). Seaweed polysaccharides as potential therapeutic agents against white spot syndrome virus (WSSV): A mini review. *Aquaculture International*, *28*, 2333–2343.

Wang, H. C., Hirono, I., Maningas, M. B. B., Somboonwiwat, K., & Stentiford, G. (2019). ICTV virus taxonomy profile: Nimaviridae. *Journal of General Virology*, *100*(7), 1053–1054. https://doi.org/10.1099/jgv.0.001248

How to cite this article: Namitha R, Vasantharaja R, Radhakrishnan M, Govindaraju K. Actinobacteria and their bioactive molecules for anti-WSSV activity: A mini review. *Aquac Res*. 2020;00:1–7. https://doi.org/10.1111/are.14942

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0

Aadhar Card Verification Base Online Polling

Ch.sai pratap varma	D.sumanth rahul	Jithina Jose	B Keerthi Samhitha	Suja Cherukullapurath Mana
Department of CSE	Department of CSE	School of Computing	School of Computing	School of Computing
Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology	Sathyabama institute of science and technology
Pratapvarma003@gmail.com	prakashnaidu073@gmail.com	jithinajose@gmail.com	samhitha711@gmail.com	cmsuja@gmail.com

Abstract - Currently, bogus voting in elections is remaining as a serious issue. The question of casting a ballot as far as health and protection are concerned is still important. The undertaking is primarily planned for giving a most verified and easy to understand Online Election System. An online political decision system is a voting system. In Traditional political race condition casting a ballot procedure now and again become very badly designed because of the reluctances of specific voters to visit a surveying stall to cast casts a ballot other than including tremendous social and HR. This paper manages the development and progress of an electronic voting system using distributed computing and aadhar card to provide the voting system with a superior with high security. Additionally, web innovation is used to make the voting system reasonable. The Internet surveying system is the system that empowers the client to cast a ballot on the web. In particular, the Voting system comprised of server administrations which are each connected with a database for putting away persevering information. Voters can likewise utilize the administrations to sign into the electronic voting system site. The field official will send the voter data and their voting's and that data will be kept up by creator card check. Administrators will keep up all data concerning voter and tallies consequently their deciding in favor offs their chosen gatherings.

Keywords: ***Electronic voting system, Web-based voting system***

I. INTRODUCTION

India is a law-based country which has a population of 700 million. The Indian Government means "all communities, people, and organizations." Each person has the right to cast a ballot and choose the creator of his / her choosing. The voting system started in India in the 18th century. The voting age for every voter is 18 years or more. In the new voting system, the security problems caused a ballot to be cast online. To that the shortcomings found in e-voting, the biometric casting of a ballot system is used. Biometric alludes to each person's unique evidence, including voice, iris, unique fingerprint, face, and palm. In the aftermath of the British government seizing the opportunity, the Indian government offers Indian individuals a right to choose their leader. In this way, the errand of political race commission is checked to achieve the straightforwardness through casting a ballot system. The object of the vote is to encourage electors to elect their legislature and political delegates. The political race is the basic voting process by which people express their feelings by selecting their opponents. India is passing huge cash to strengthen our voting system so that people can have a superior government. The voting system should be open, versatile, and entirely secure to ensure a better voting system. The present scheme is used to make it less obvious because a ballot period for the equipment can be cast. Voting process confidentiality, authentication of voters, ensuring the cast ballot details are the major variables of the current political decision system.

Authorized licensed use limited to: Auckland University of Technology. Downloaded on August 09,2020 at 08:57:59 UTC from IEEE Xplore. Restrictions apply.

The object of casting a ballot is to enable electors to exercise their parliamentary and political right to select their representatives. Innovation is increasingly being used as a method to help electors cast their ballots. Practically all voting systems across the globe require the following technologies to allow this right to operate: Voter identification and authentication, Vote and register cast ballots, Vote verification, Reporting democratic election results. Throughout two periods of the appointment process, the voter-identifiable proof is required: first, for voter registration to obtain the privilege of casting a ballot, and shortly afterward, when casting a ballot, enabling a citizen to exercise his right to cast a ballot by ensuring that the individual meets all the anticipated conditions of casting a ballot[9].

This calculation is strong, precise, and coordinating calculation for most common specific finger impressions. The image is updated for further preparation and revamped to the parallel image. In the second step, the picture is diminished and the details are removed. Ultimately, information sets of two fingerprints are synchronized to obtain a coordinated score. This technique will detect the similarity between the test of input information and the test of putting away info without using the fatiguing test[10]. The cycle of the political race takes on significant work within a just country. The Election is a process of choosing an ideal candidate to lead the country.

The electronic voting system is used from late in India. Accessibility is needed right now in the region; this is a major downside to the electronic voting system. The system is an electronic voting scheme since an elector from anywhere will cast a vote. The new AADHAR card online voting system allows voters to test their fingerprints, which is then compared to an image previously spared inside a database obtained from the AADHAR card database. A poll clerk casts a hidden key at the time to diminish fraudulent voting. Survey software is the platform that now empowers clients to cast a ballot on the internet. The Voting system consisted in particular of server administrations, each of which is linked to a database for storing tenacious information. Voters may also use the administrations to sign onto the site of the electronic voting system. The administrator shall keep all data concerning the voter and shall, therefore, verify that they agree in favor of their chosen meetings.

II. RELATED WORK

Uses PIN, biometric photos, and steganography for authentication in Secure Online Voting System Authentication. Hash code is used to verify whether or not steganography image is being altered. Calculating the number of pixels used would make the recovery of PIN impossible [1]. Effectively configured using Fingerprint and Hex keypad in the AADHAR-based Electronic Voting System. The server composed of the individual subtleties should be modified each time before the election. A perfect E-Voting system which allows impeccable safety and protection without compromise [2] is extremely difficult to structure.

In Fingerprint and RFID Based Electronic Voting System Linked with AADHAR, a Biometric voting system based on RFID security is provided. Furthermore, it offers protection from the drunken individual who arrives at the survey corner. Important special features such as Voter ID, AADHAR ID, and Biometric Validation have been identified with this program. In Election, there is no degree of ragging[3]. The introduction of the Authenticated and Secure Online Voting Program helps achieve genuineness, non-discernibility of casting votes, and anonymity with the same authorization. It gives a secure secret ballot at the enrollment hour which empowers the voter to make their choice safely alongside biometric distinguishing evidence [4].

The polling officer will press the Ballot button instead of issuing a form of a ballot vote. Instead of sending papers on a ballot sheet, polling officials press the button on the voting sheet. This will allow the elector to make his or her choice by using methods to pinch the blue catch against the competitors and the picture of their panel decisions. [6]. [18]. There are a few significant advances to the Internet casting a ballot system. The structure is congenial from various sides: (I) from the perspective of India's election commission, which is also the administrator. (ii)From voting centers. Which allows electors to cast the votes on E-Ballot paper [1].

Authorized licensed use limited to: Auckland University of Technology. Downloaded on August 09,2020 at 08:57:59 UTC from IEEE Xplore. Restrictions apply.

III. EXISTING SYSTEM

To dispose of the new system's drawbacks, A program has been developed whereby the user does not need to spend a lot of energy in the site test, Or rather server is manufactured that deals with the only voting procedure as customer needs to offer his voter Id along those lines decreasing the weight on the voters and the official sector.

3.1 DISADVANTAGES OF THE EXISTING SYSTEM

3.1.1 EXPENSIVE AND TIME DEVOURING

The way that information is obtained and stored in the database takes a long time and is costly to lead, for example, To the time and money spent in the printing of information collection systems, in the planning of registration stations along with HR, and there in the wake of the promotion of the days set for the registration process, Remembering the sharpening of the identification criteria for voters, even as time was spent on incorporating this data into the database.

3.1.2 TOO MUCH ADMINISTRATIVE WORK

The method requires a lot of office work and paper storage which is troubling because papers are big with population size.

IV. PROPOSED SYSTEM

In the proposed system, Aadhar Database is used as the backbone of our project to verify and test voting qualifications. This determines by its special finger impression on the actual voter whether or not he/she is a significant voter. This helps the individual voter to make the choice on the web and to update the server database. The biometric web-based voting system uses the Aadhar card to retrieve absolute information about the voter. The proposed model has more notable protection as the high-security hidden phrase is confirmed before the vote is recognized in the Election Commission of India's key database. The voting system involves the physical and managerial partition of the appointive register and electronic polling booth. In particular, the Casting a ballot system comprised of server administrations which are each connected with a database for creator card confirmation.

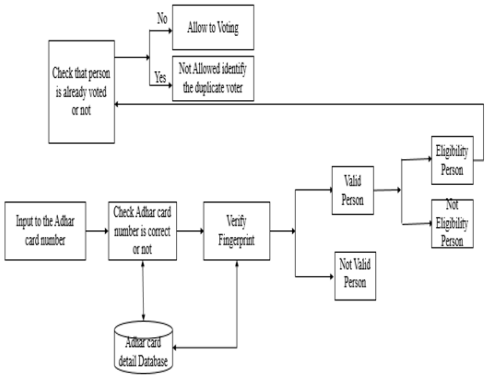


Fig.1

4.1 ADVANTAGES OF THE PROPOSED SYSTEM

- Saved Ballot Models have the requirement to decide without preparation. Do what needs to be done once, at that point spare that voting form design, and in ensuing years, indicate just the names of the applicants.
- Minimized costs are met when the cost of printing, shipping, and arranging paper polling forms is minimized or even completely dispensed from the political race process.

V. MODULE DESCRIPTION

5.1 COMPONENTS OF THE SYSTEM

The customer server web-empowered programming engineering for the undertaking. A special mark scanner is used on the customer side and a GUI that recognizes the aadhar number of the voter provides an interface to cast a ballot and display approval, rank, and blunder message. The GUI will only follow up on occasions with no additional planning from the server and voter feedback. Servers are set out from the survey corners at remote locations. All of the handling works can be completed, for example, picture preparing, moving information between the customer and the database, producing insights, sending messages to voters, and so on. There is a focal database called the Central Data Repository of Identities, or CIDR, which forms the basis of this structure.

Authorized licensed use limited to: Auckland University of Technology. Downloaded on August 09,2020 at 08:57:59 UTC from IEEE Xplore. Restrictions apply.

5.2 AUTHENTICATION AND REQUIREMENT OF THE VOTER

Confirmation is the way to determine whether a person or object is being told fact, who or what it is being declared to be. It has been expected to have a valid UID number to validate a person. It is assumed that they will have a valid UID number for validating an individual. The number will be checked in the nearby database records first. On the off chance that it isn't discovered, at that point, it will look through the focal storehouse. It includes a one-to-many match. On the off chance that the individual's number isn't found in the focal database, at that point obviously s/he will be without participating in the voting procedure. And again if the number is available in the focal database, it will store the individual's information in the zonal database at that level. This record is deleted from the database in the neighborhood and sent to servers to be reviewed for further planning. To check the specific finger impression of the person on the customer side will be filtered and balanced synchronized on the servers with different information from the nearby database. This method places less weight on the community database and improves the information about traffic. For authentication, fingerprints are used because fingerprint preparation is faster and superior to other biometric details. Besides aadhar subtleties are lacking to establish the true character of a person as they can be faked effectively but by using fingerprints it is guaranteed that such fake parts will be blocked right at the earliest point of reference.

5.3 PREVENTING FRAUDULENT VOTING

The first and foremost thing to ensure adequate voting is to precise approval of each vote. It is important to understand that each person who comes to cast a ballot is one of a kind else it will destroy the very principle of casting a ballot. Any individual would decide for the benefit of others. A unique coordinating mark guarantees the confirmation needed by the device. In any case, it is necessary to keep the bogus rejection rate (FRR) and the bogus acceptance rate (FAR) as small as possible to improve accuracy; for all intents and purposes, close to zero. Using the best finger technique will achieve this. Using 2 best fingers will increase the accuracy with a single effort to 98 percent (FRR-2 percent) or more than 99 percent (FRR–1 percent) with up to 3 devours. To prevent underage people from casting a vote, the program displays the age of the person in the database records from the date of birth. If the determined age is above passable cutoff the individual is permitted to cast a ballot and forestalled in any case. To keep voters from casting a ballot two or on different occasions that can actualize casting ballot signals in the neighborhood databases. This banner is at first set to bogus. His / her vote banner will be set to genuine after the elector has made the option. It will prohibit a similar individual from casting a ballot once again.

These banners are impermanent and can be reset after the political race has finished so the voter can take an interest in the following political decision. Giving straightforwardness is another type of keeping acts of neglect in casting a ballot procedure from a voter's perspective. Since a voter needs to place his trust into PC's equipment and programming those procedures the vote, it doesn't ensure straightforwardness in the most genuine sense. Along these lines, an answer can be given to manufacture the voter's trust in the system by making an impression on the voters enrolled versatile expressing that s/he has effectively cast a ballot with the comparing aadhar number at a predetermined time.

5.4 GENERATING REPORTS

At whatever point a voter chooses the up-and-comer of decision, the vote tally of that up-and-comer gets augmented in the neighborhood database. Voting from all party sources is weighted to get the final number the participant has won. Through this, this system will yield immediate results and avoid inefficient resource use and waste time. As this is an advanced computing device that uses the knowledge it has a few points of interest. Insights can be created from the acquired information, for example, we could answer what number of individuals have cast a ballot from a specific district, what number of females cast a ballot, which age bunch cast a ballot the most, the most elevated turnouts, correlations from earlier years, and so forth every one of that was unrealistic from conventional voting strategies not even from

Authorized licensed use limited to: Auckland University of Technology. Downloaded on August 09,2020 at 08:57:59 UTC from IEEE Xplore. Restrictions apply.

EVMs. It would give significant bits of knowledge into the political decision results and help improve the system considerably further.

VI. CONCLUSION

This voting system unquestionably lets everyone cast their ballots. Web-based voting should create the degree of ballot casting. It doesn't need manual tallying. So we're going to get the incredibly obvious, transparent and quick result. By using this newly created system can solve numerous problems of the existing system. By using the aadhar token, the modified framework extends the protection of the voter. The system supervised a ballot program in less complex ways just as safe to cast. The aadhar will give each voter the special ID so that the security break is kept away. This approach is more successful than the present one. A web-based voting system is developed in a better and quicker way than the previous systems, which ensure transparency, continuity, and retains upright voting procedures. It additionally permits an individual to cast a ballot from anyplace given that the voter is inside discretionary points of confinement.

REFERENCES

[1]. Smita B. Khairnar,P. Sanyasi Naidu, ReenaKharat ,"Secure Authentication for Online Voting System" International Journal of Computer Science and Information, 2015

[2]. Prof. R. L. Gayle, Vishnu Lokhande, Shubham T. Jaddhav, "Aadhar Based Electronic Voting System" International Journal of Advance Scientific Research and Engineering Trends, May 2016

[3]. B. Mary Haque, G. M. owais Ahmed, "Fingerprint and RFID Based Electronic Voting System Linked with Aadhar for Rigging Free Election", International Journal of Advance Research in Electrical, Electroinc and Instrumentation Engineering, March 2016

[4]. Srivatsan Sridharan, "Implementation of Authenticated and Secure Online Voting System",IEEE,4th july 2013

[1] H.Agarwal and G.N.Pandey, "Online Voting System For India Based On Aadhar Id," 2013 Eleventh International Conference on ICT and Knowledge Engineering, Bankok,2013.

[2] Gowri,Guruprasanth, Jaya Surya. D, Krishnan. S, Dhanasekaran.S.,"Imphilementation of Biometric Voting Machine Using Aadhar Card," IJSRSET, Volume 2, Issue 2, 2016.

[3] G.Valarmathy, V.Saranya, Riya Rose Cherian, K.Poovizhi,"Smart Voting System Using Aadhar", published in IJRE, Volume: 04 Issue: 01 2017.

[4] Ankita R. Kasliwal, Jaya S. Gadekar. Manjiri A. Lavadkar, Pallavi K. Thorat, Dr. PraptiDeshmukh., " Aadhar Based Election Voting System", IOSR Journal of computer engineering.

[5] Anandaraj S, Anish R, Devakumar P.V, "Secured Electronic Voting Machine Using Biometric", ICIIECS 2015.

[6] Murali R Prasad, PolaiahBojja, and MadhuNakirekanti, "Aadhar Based Electronic Voting Machine Using Arduino", International Journal of Computer Applications 145(12):39-42, July 2016.

[7] SoumyajitChakraborty, Siddhartha Mukherjee, BhaswatiSadhukhan, KaziTanviYasmin,"Biometric Voting System Using Adhar Card in India", Published in IJIRCCE Vol. 4, Issue 4, April 2016.

[8] SaurabhYadav and Ajay Kr. Singh, "A Biometric Traits based Authentication System for Indian Voting System", International Journal of Computer Applications (0975 – 8887) Volume 65– No.15, March 2013

[9] Prof.R.L.Gaike, Vishnu P. Lokhande, Shubham T. Jadhav and Prasad N. Paulbudhe, "Aadhar Based Electronic Voting System", Volume 1, Issue 2 May 2016, ISSN (Online) 2456-0774 International Journal Of Advance Scientific Research and Engineering Trends.

[10] W. Chen and Y. Gao, "A Minutiae-Based Fingerprint Matching Algorithm Using Phase Correlation," 9th Biennial Conference of the Australian PatternRecognition Society on Digital Image Computing Techniques and Applications (DICTA 2007), Glenelg, Australia, 2007, pp. 233-238.doi:10.1109/DICTA.2007.4426801

[11] J.Jose and T Sasipraba *"Indoor air quality monitors using IOT sensors and LPWAN",2019* 3rd International conference on Trends in electronics and informatics(ICOEI), Tirunelveli, India, 2019, PP 633-637.

[12] "*An Efficient Implimentation of Methods to Detect Sybil Attacks in Vehicular Ad hoc Networks Using Received Signal Strength Indicators* ",B.Keerthi Samhitha , SujaCherukullapurathMana **, Jithina Jose**, M. Mohith, L. Siva Chandhrahasa Reddy
International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November, 2019

[13] "*An Efficient System to Predict and Analyze Stock Data using Hadoop Techniques*", **Jithina Jose**, SujaCherukullapurathMana, B.Keerthi Samhitha International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2, July 2019

[14] *"Traffic violation detection using principal component analysis and viola jones algorithms"*, SujaCherukullapurathMana, **Jithina Jose,** B.Keerthi Samhitha International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Authorized licensed use limited to: Auckland University of Technology. Downloaded on August 09,2020 at 08:57:59 UTC from IEEE Xplore. Restrictions apply.

A Unique Extensible Framework Detection and Classification of Traffic Events Based on Deep Learning Approach

Prudhvikrishna Gangavarapu, Sandeep Gunapaneni & K. Mohana Prasad

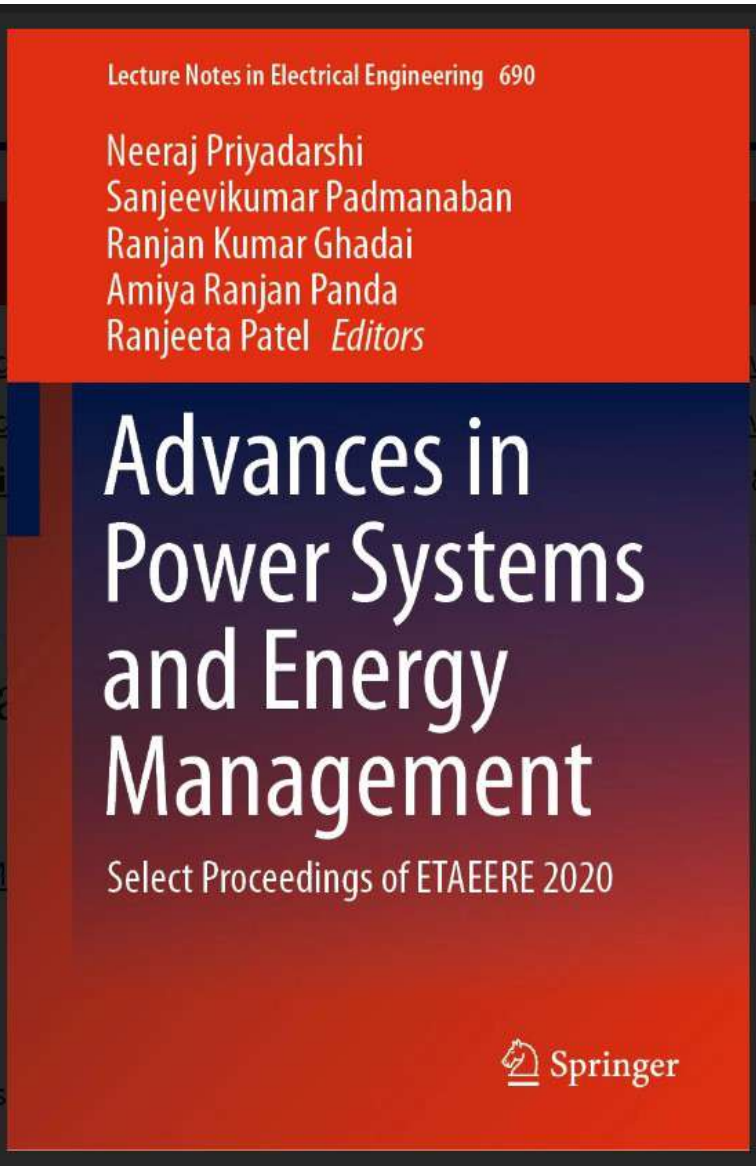
Conference paper | First Online: 21 January 2021

307 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 690)

Abstract

Social media, like Twitter, Instagram, YouTube, Facebook, Tumblr etc. is used to obtain valuable information relevant to traffic, which may be a commentary or user produced encounter. In this skeptical representation we mainly use Deep belief network (DBN) which was one of the major Deep Learning Techniques that will be utilized for public sentiment evaluation on traffic information through social networking. In the current system, the visitor's information may be decided to work with Google Maps or perhaps in FM Radio. Google Map won't provide the correct Traffic Information. Additionally, it won't supply some info regarding


Lecture Notes in Electrical Engineering 690
Neeraj Priyadarshi
Sanjeevikumar Padmanaban
Ranjan Kumar Ghadai
Amiya Ranjan Panda
Ranjeeta Patel Editors
Advances in Power Systems and Energy Management
Select Proceedings of ETAEERE 2020
Springer

Cite this paper

Gangavarapu, P., Gunapaneni, S., Prasad, K.M. (2021). A Unique Extensible Framework Detection and Classification of Traffic Events Based on Deep Learning Approach. In: Priyadarshi, N., Padmanaban, S., Ghadai, R.K., Panda, A.R., Patel, R. (eds) Advances in Power Systems and Energy Management. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 690. Springer, Singapore. https://doi.org/10.1007/978-981-15-7504-4_58

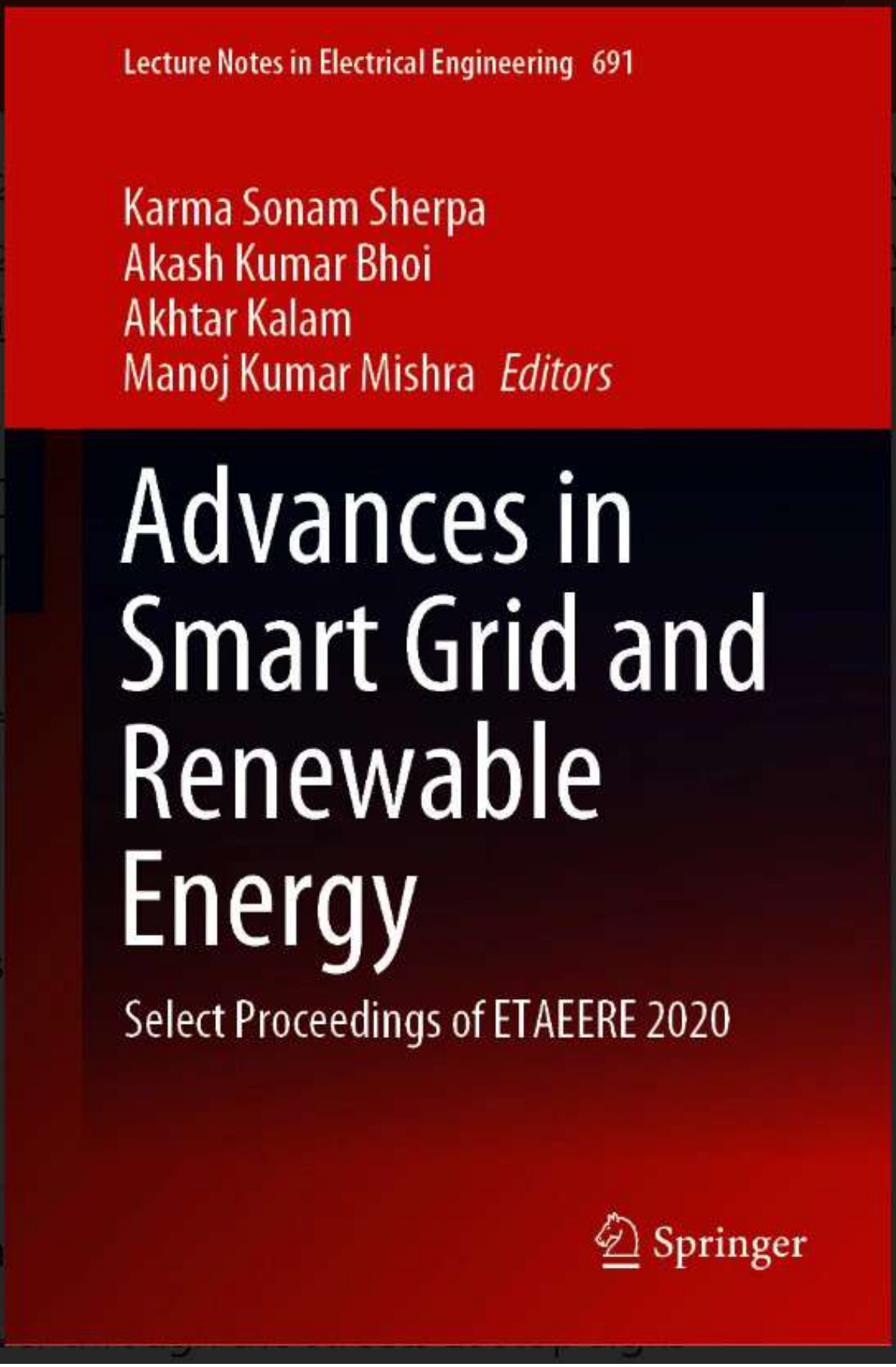
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7504-4_58

Published	Publisher Name	Print ISBN
21 January 2021	Springer, Singapore	978-981-15-7503-7
Online ISBN	eBook Packages	
978-981-15-7504-4	Energy Energy (R0)	


Lecture Notes in Electrical Engineering 691
Karma Sonam Sherpa
Akash Kumar Bhoi
Akhtar Kalam
Manoj Kumar Mishra Editors
Advances in Smart Grid and Renewable Energy
Select Proceedings of ETAEERE 2020
Springer

A Unique Adaptive Framework for Predicting Lane Changing Intention Based on CNN

K. Pavan, M. Dhanaveera Teja, A. Pravin, T. Prem Jacob & G. Nagarajan

Conference paper | First Online: 05 January 2021

315 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 691)

Abstract

To order for vehicles to be able to drive on their own, they need to understand the world around them as human drivers, so that they can maneuver through the streets at stop signs and traffic lights to avoid hitting obstacles such as other cars and pedestrians. Based on the problems encountered when autonomous vehicles detected objects, an effort was made to display lane detection using OpenCV library. The reason and procedure for selecting grayscale rather than color, detecting the edges in an image, selecting region of interest, transforming Hough and selecting polar coordinates over Cartesian coordinates were discussed. For this various development done on autonomous cars, and evaluation is done based on tools

Cite this paper

Pavan, K., Teja, M.D., Pravin, A., Jacob, T.P., Nagarajan, G. (2021). A Unique Adaptive Framework for Predicting Lane Changing Intention Based on CNN. In: Sherpa, K.S., Bhoi, A.K., Kalam, A., Mishra, M.K. (eds) Advances in Smart Grid and Renewable Energy. ETAEERE ETAEERE 2020 2020. Lecture Notes in Electrical Engineering, vol 691. Springer, Singapore. https://doi.org/10.1007/978-981-15-7511-2_58

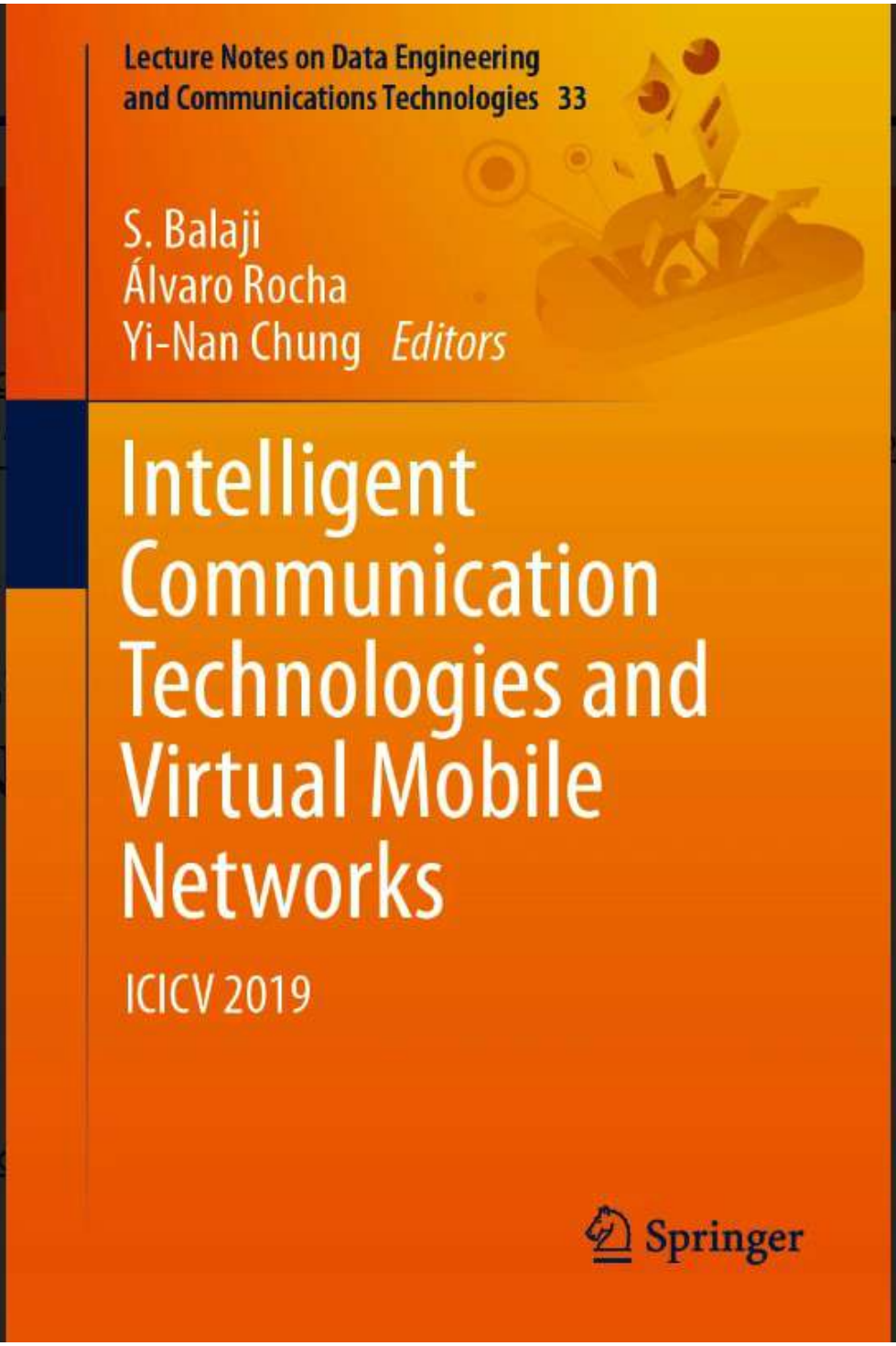
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7511-2_58

Published	Publisher Name	Print ISBN
05 January 2021	Springer, Singapore	978-981-15-7510-5
Online ISBN	eBook Packages	
978-981-15-7511-2	Energy Energy (R0)	


Lecture Notes on Data Engineering and Communications Technologies 33
S. Balaji
Álvaro Rocha
Yi-Nan Chung Editors
Intelligent Communication Technologies and Virtual Mobile Networks
ICICV 2019
Springer

A Study of Biology-Based Congestion Control Algorithms for Wireless Sensor Network

S. Panimalar & T. Prem Jacob

Conference paper | First Online: 13 August 2019

864 Accesses

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 33)

Abstract

Network Traffic is one of the major issues in wireless Sensor Networks (WSNs). WSN is a self-constructed and organization less wireless networks which is used to observe and check the physical or environmental conditions and to cooperatively pass their data through the network to a sink where the data can be appropriately observed and examined. Number of research works in wireless sensor networks (WSNs) is primarily focused on improving the network performance along with enhancing the quality of service parameters such as the data arrival rate, available bandwidth, congestion, transmission rate, queue length and energy.

Cite this paper

Panimalar, S., Prem Jacob, T. (2020). A Study of Biology-Based Congestion Control Algorithms for Wireless Sensor Network. In: Balaji, S., Rocha, Á., Chung, YN. (eds) Intelligent Communication Technologies and Virtual Mobile Networks. ICICV 2019. Lecture Notes on Data Engineering and Communications Technologies, vol 33. Springer, Cham. https://doi.org/10.1007/978-3-030-28364-3_25

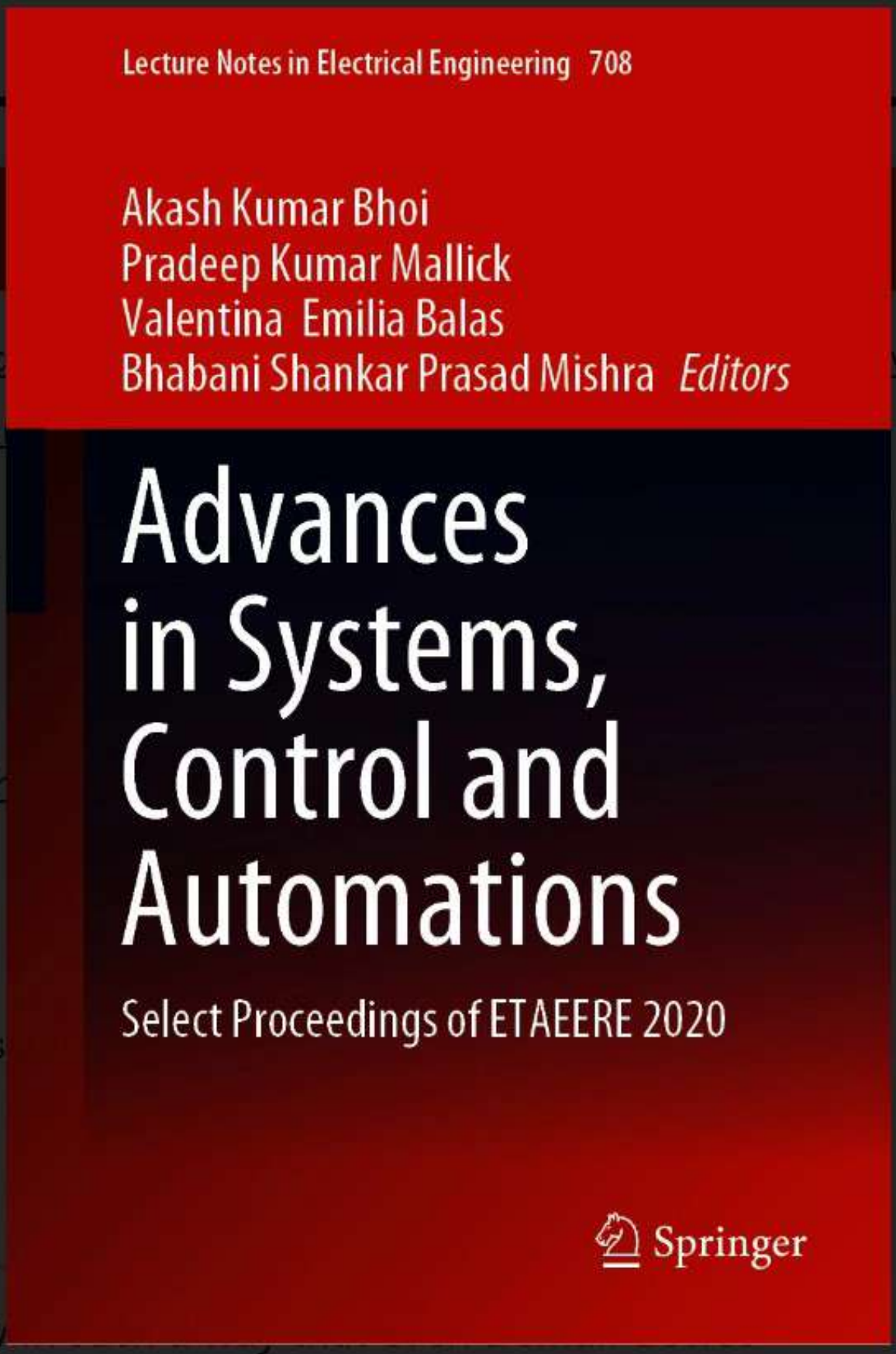
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-28364-3_25

Published	Publisher Name	Print ISBN
13 August 2019	Springer, Cham	978-3-030-28363-6
Online ISBN	eBook Packages	
978-3-030-28364-3	Engineering Engineering (R0)	


Lecture Notes in Electrical Engineering 708
Akash Kumar Bhoi
Pradeep Kumar Mallick
Valentina Emilia Balas
Bhabani Shankar Prasad Mishra Editors
Advances in Systems, Control and Automations
Select Proceedings of ETAEERE 2020
Springer

Cite this paper

Bhuvana, K., Bhargavi, B.K., Vigneshwari, S. (2021). A Safety Stick for Elders. In: Bhoi, A.K., Mallick, P.K., Balas, V.E., Mishra, B.S.P. (eds) Advances in Systems, Control and Automations . ETAEERE 2020. Lecture Notes in Electrical Engineering, vol 708. Springer, Singapore. https://doi.org/10.1007/978-981-15-8685-9_53

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-8685-9_53

Published	Publisher Name	Print ISBN
20 March 2021	Springer, Singapore	978-981-15-8684-2
Online ISBN	eBook Packages	
978-981-15-8685-9	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)	

A Safety Stick for Elders

Korrapati Bhuvana, Bodavula Krishna Bhargavi ✉ & S. Vigneshwari

Conference paper | First Online: 20 March 2021

199 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 708)

Abstract

Technology in recent decades is developing day-by-day in such a way that even a small device is equipped with many features. In this paper, we provide a detailed prototype of a safety stick that helps the elders to lead their life independently. Most commonly elders use a walking stick for their support, we are going to make that normal walking stick into a smart safety stick. Here, we use IoT and embedded technology. In an emergency like when they need medical help, by long-pressing, a simple button implemented walking stick the alert message is sent to the caregiver along with location information. Also, when in the threat of thieves they can use the same stick for their protection by pressing simple click it will give a shock to the thief and also a message is sent to caregiver simultaneously and even the neighbors with a


S. Smys
Abdullah M. Iliyasu
Robert Bestak
Fuqian Shi Editors
New Trends in Computational Vision and Bio-inspired Computing
Selected works presented at the ICCVBIC 2018, Coimbatore, India
Springer

A Review on False Data Injection in Smart Grids and the Techniques to Resolve Them

P. Asha, K. Deepika, J. Keerthana & B. Ankayarkanni

Chapter | First Online: 28 September 2020

38 Accesses

Abstract

Smart grid plays a vital role in electricity supply network by using digital communication technology. It is capable of easily detecting and reacting to the local changes in the usage and thereby working efficiently. Against all the smart grid monitoring system there are many cyber intrusion which creates severe threat to power to power system operation and false data injection is one among them. In this paper we discuss about the different techniques. Kalman filter, state estimation, bad data detection test, Phasor measurement units (PMU) installation are used to detect and minimize false data injection attack.

Cite this chapter

Asha, P., Deepika, K., Keerthana, J., Ankayarkanni, B. (2020). A Review on False Data Injection in Smart Grids and the Techniques to Resolve Them. In: Smys, S., Iliyasu, A.M., Bestak, R., Shi, F. (eds) New Trends in Computational Vision and Bio-inspired Computing. ICCVBIC 2018. Springer, Cham. https://doi.org/10.1007/978-3-030-41862-5_153

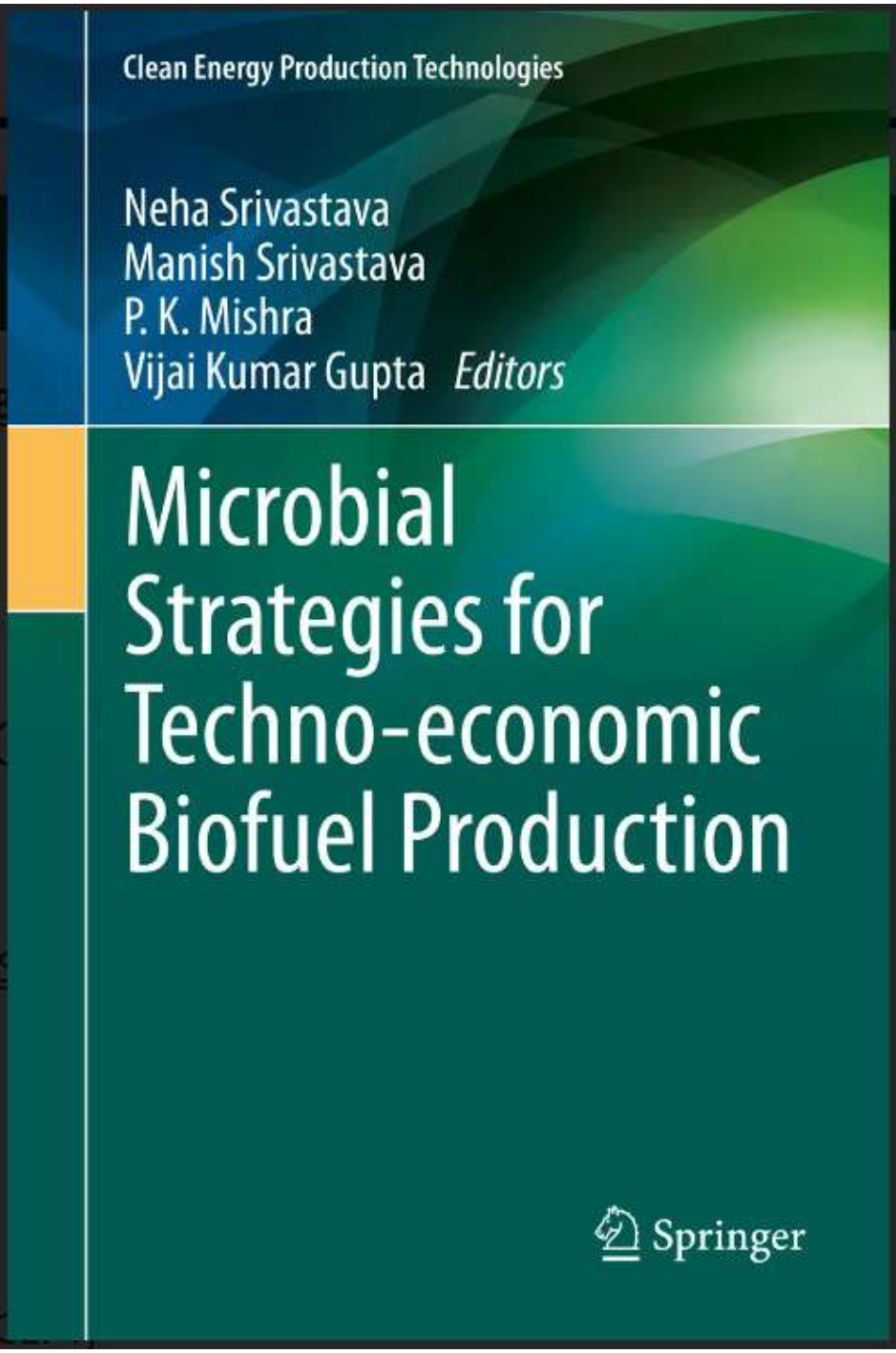
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-41862-5_153

Published	Publisher Name	Print ISBN
28 September 2020	Springer, Cham	978-3-030-41861-8
Online ISBN	eBook Packages	
978-3-030-41862-5	Mathematics and Statistics Mathematics and Statistics (R0)	


Clean Energy Production Technologies
Neha Srivastava
Manish Srivastava
P. K. Mishra
Vijai Kumar Gupta Editors
Microbial Strategies for Techno-economic Biofuel Production
Springer

Nanocatalysts to Improve the Production of Microbial Fuel Applications

Siva Sankar Sana, Punita Kumari, Zeba Usmani, Minaxi Sharma, Surya Sudheer, D. Dinesh Kumar, Zhijun Zhang & Huizhen Li

Chapter | First Online: 13 September 2020

211 Accesses

Part of the Clean Energy Production Technologies book series (CEPT)

Abstract

Fossil fuel which is the current source for energy supply is responsible for environmental pollution. Recently, renewable source comes as a substitute for this because of the depletion of fossil fuel resources and thus its inability to meet the future energy demand. Renewable energy now used is the wind, solar, biomasses, hydrogen and geothermal source. Among them, biomass is more ideal because of its efficiency in converting to liquid biofuel directly.

Cite this chapter

Sana, S.S. *et al.* (2020). Nanocatalysts to Improve the Production of Microbial Fuel Applications. In: Srivastava, N., Srivastava, M., Mishra, P.K., Gupta, V.K. (eds) Microbial Strategies for Techno-economic Biofuel Production. Clean Energy Production Technologies. Springer, Singapore. https://doi.org/10.1007/978-981-15-7190-9_8

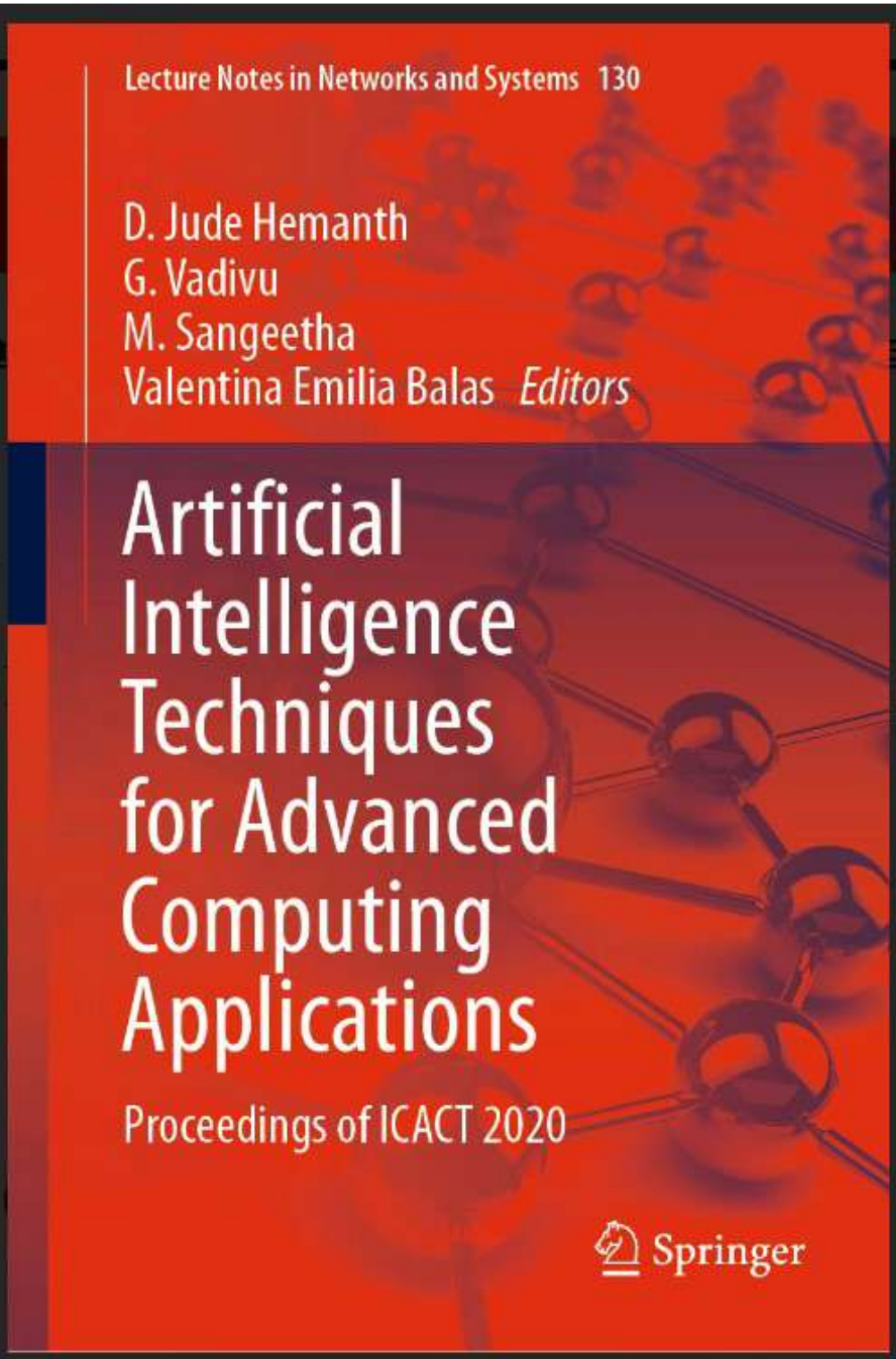
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-7190-9_8

Published	Publisher Name	Print ISBN
13 September 2020	Springer, Singapore	978-981-15-7189-3
Online ISBN	eBook Packages	
978-981-15-7190-9	Biomedical and Life Sciences Biomedical and Life Sciences (R0)	


Lecture Notes in Networks and Systems 130
D. Jude Hemanth
G. Vadivu
M. Sangeetha
Valentina Emilia Balas Editors
Artificial Intelligence Techniques for Advanced Computing Applications
Proceedings of ICACT 2020
Springer

A Comprehensive Study on Ransomware Attacks in Online Pharmacy Community

V. Joseph Raymond & R. Jeberson Retna Raj

Conference paper | First Online: 24 July 2020

478 Accesses

Part of the Lecture Notes in Networks and Systems book series (LNNS,volume 130)

Abstract

In recent trends, there are many cyber-attacks done across various organizations around the globe. The largest pharmaceutical companies found to be comprised of a ransomware Petya malicious virus. Online pharmacy communities are doing their best for delivering care and treatment with Internet support, but the attackers exploit vulnerabilities in operating systems to identify software flaws. The attackers access patient summary records and threaten pharmacy owners to pay a substantial sum to prevent leaking confidential information and henceforth locking the system by sending phishing e-mails across network and seed

Cite this paper

Joseph Raymond, V., Retna Raj, R.J. (2021). A Comprehensive Study on Ransomware Attacks in Online Pharmacy Community. In: Hemanth, D., Vadivu, G., Sangeetha, M., Balas, V. (eds) Artificial Intelligence Techniques for Advanced Computing Applications. Lecture Notes in Networks and Systems, vol 130. Springer, Singapore.
https://doi.org/10.1007/978-981-15-5329-5_6

Download citation

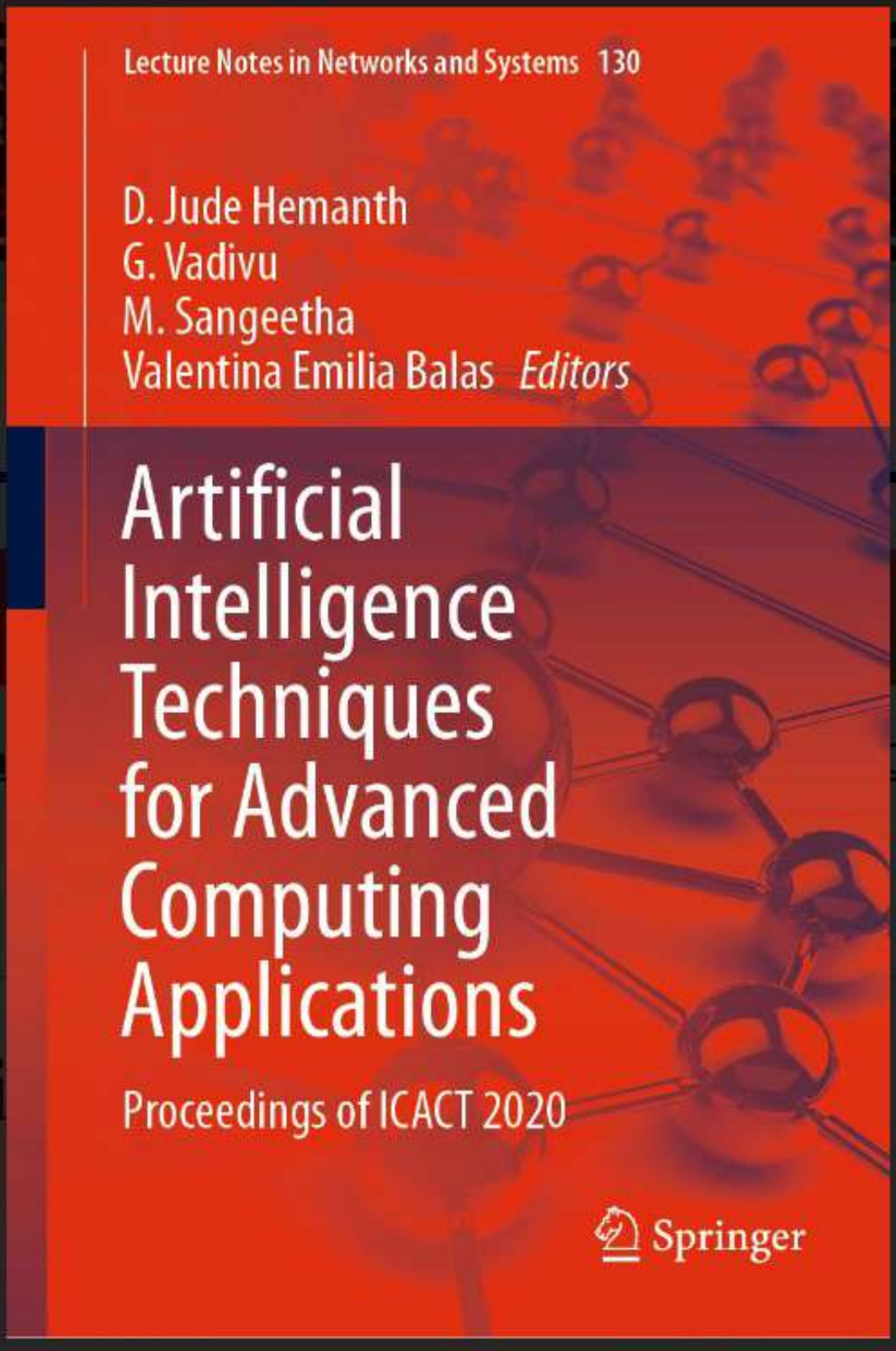
.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-5329-5_6

Published	Publisher Name	Print ISBN
24 July 2020	Springer, Singapore	978-981-15-5328-8
Online ISBN	eBook Packages	
978-981-15-5329-5	Intelligent Technologies and Robotics	

Advertisement



A Study on Image Hashing Techniques for Implementing Deduplication

G. Sujatha & Jeberson Retna Raj

Conference paper | First Online: 24 July 2020

518 Accesses | 1 Citations

Part of the Lecture Notes in Networks and Systems book series (LNNS,volume 130)

Abstract

Image deduplication is an approach to avoid duplicate images to be stored in cloud, which is currently required to increase the efficient utilization of cloud storage. Most of the times, the images uploaded by different users of cloud storage service need not to be unique. There is a possibility of storing same images by different users. This may lead to poor usage of cloud storage. To find duplicates, various image hashing techniques are widely used. The purpose of this paper is to review various image hashing methods which are used to generate hash value of the images. Using the hash value, we can conclude the uniqueness of the image. This can

Cite this paper

Sujatha, G., Raj, J.R. (2021). A Study on Image Hashing Techniques for Implementing Deduplication. In: Hemanth, D., Vadivu, G., Sangeetha, M., Balas, V. (eds) Artificial Intelligence Techniques for Advanced Computing Applications. Lecture Notes in Networks and Systems, vol 130. Springer, Singapore. https://doi.org/10.1007/978-981-15-5329-5_49

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-15-5329-5_49

Published	Publisher Name	Print ISBN
24 July 2020	Springer, Singapore	978-981-15-5328-8
Online ISBN	eBook Packages	
978-981-15-5329-5	Intelligent Technologies and Robotics	

Lecture Notes on Data Engineering
and Communications Technologies 26



Jude Hemanth
Xavier Fernando
Pavel Lafata
Zubair Baig *Editors*

International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018

Springer

User Based Efficient Video Recommendation System

J. Albert Mayan, R. Aroul Canesaane, J. Jabez, M. D. Kamalesh & G. Rama Mohan Reddy

Conference paper | First Online: 21 December 2018

1357 Accesses | 2 Citations

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 26)

Abstract

The quick increment in the online data content has made it extremely troublesome for individuals to discover data that is pertinent to their requirements and interests. Proposal framework is an intense apparatus that gives a potential answer for this test by offering a mechanized component to search out applicable and new data. Individuals depend on client surveys on the web when they require data about item, motion picture, video, news, eatery, and so on. In any case, fundamental issue is giving suggestion to client shape huge information environment. Huge information is a rising innovation, as the name recommends

Cite this paper

Albert Mayan, J., Aroul Canesaane, R., Jabez, J., Kamalesh, M.D., Rama Mohan Reddy, G. (2019). User Based Efficient Video Recommendation System. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (eds) International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018. ICICI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-030-03146-6_153

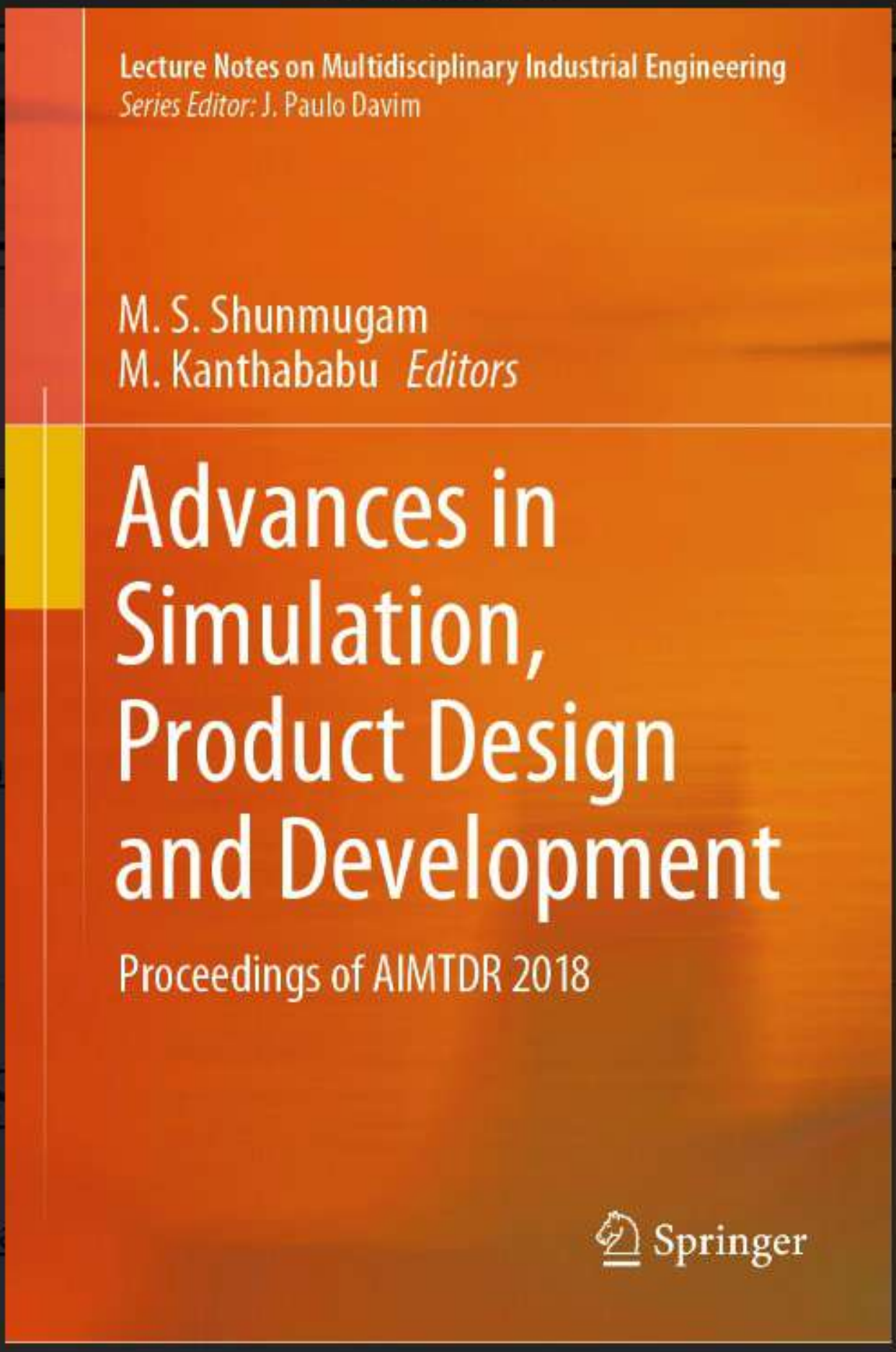
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-03146-6_153

Published	Publisher Name	Print ISBN
21 December 2018	Springer, Cham	978-3-030-03145-9
Online ISBN	eBook Packages	
978-3-030-03146-6	Intelligent Technologies and	


Lecture Notes on Multidisciplinary Industrial Engineering
Series Editor: J. Paulo Davim
M. S. Shunmugam
M. Kanthababu Editors
Advances in Simulation, Product Design and Development
Proceedings of AIMTDR 2018
Springer

Six Sigma in Battery Assembly of Skid-Steer Loader

R. Kaja Bantha Navas, S. Prakash, M. Mithun & Abhishekshivram

Conference paper | First Online: 07 November 2019

1012 Accesses

Part of the Lecture Notes on Multidisciplinary Industrial Engineering book series (LNMUINEN)

Abstract

Lead–acid battery is the vital part of any automobile, which is used to start the engine, and powers the electric controller and accessories. Quality of the battery determines the smooth starting of engine and working of electric accessories in the vehicles. A damaged battery due to wrong assembly method leads to reduced battery life and frequent starting problems in the working field. This paper eliminates cost of poor quality involved in the assembly of battery on the machine by eliminating the rejection, rework, battery surface damages, and operator fatigue by using Six Sigma methodology with improved quality by using an effective battery mounting bracket design.

Cite this paper

Kaja Bantha Navas, R., Prakash, S., Mithun, M., Abhishekshivram (2020). Six Sigma in Battery Assembly of Skid-Steer Loader. In: Shunmugam, M., Kanthababu, M. (eds) Advances in Simulation, Product Design and Development. Lecture Notes on Multidisciplinary Industrial Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-32-9487-5_56

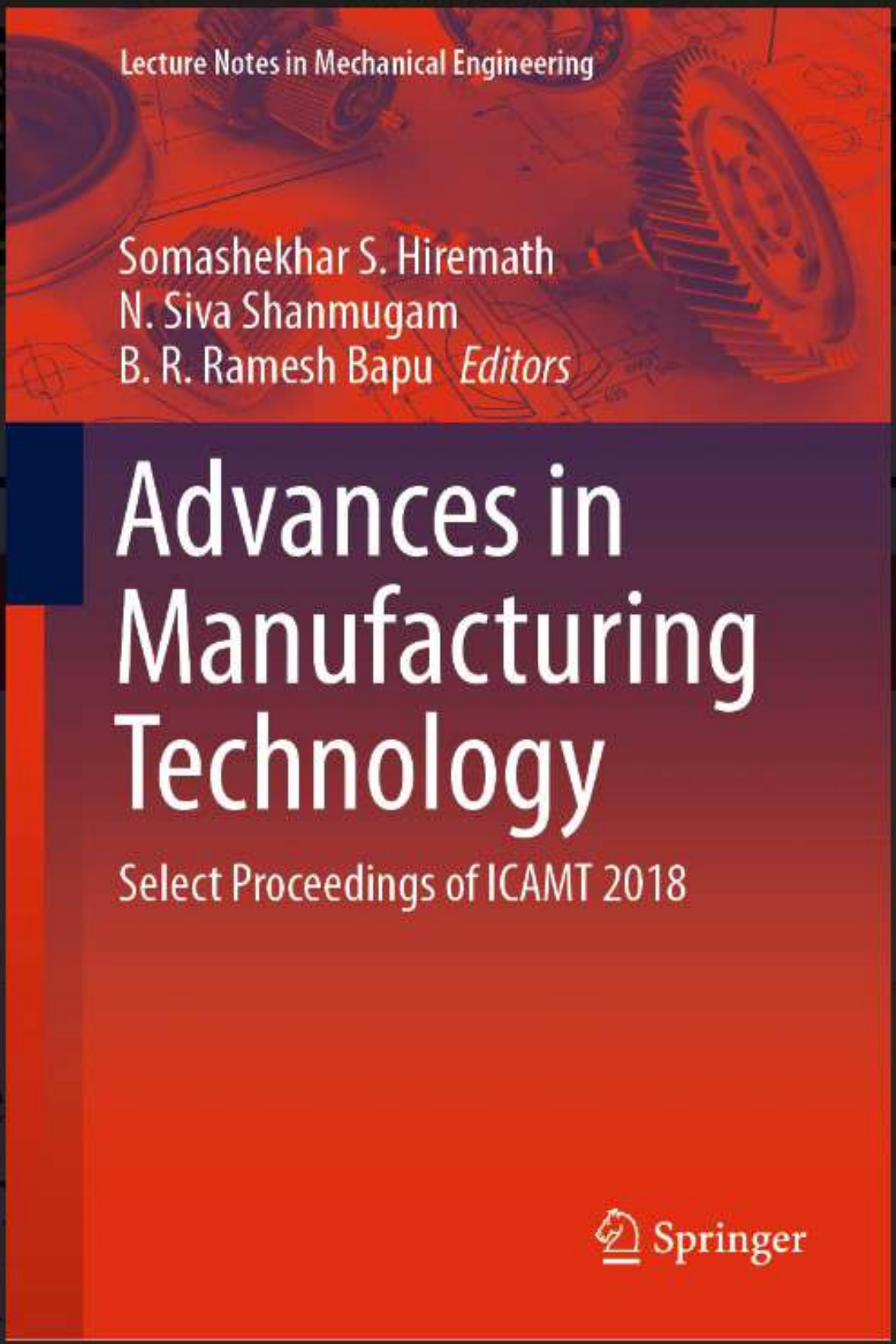
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-32-9487-5_56

Published	Publisher Name	Print ISBN
07 November 2019	Springer, Singapore	978-981-32-9486-8
Online ISBN	eBook Packages	
978-981-32-9487-5	Engineering Engineering (R0)	


Lecture Notes in Mechanical Engineering
Somashekhar S. Hiremath
N. Siva Shanmugam
B. R. Ramesh Bapu Editors
Advances in Manufacturing Technology
Select Proceedings of ICAMT 2018
Springer

Parametric Optimization of Single-Cylinder Four-Stroke Spark Ignition Engine Crankshaft Using Four Different Materials—AISI 1040, AISI 4340, AISI 4140 and AISI 4615

A. Pandiyan, G. Arun Kumar, S. Shaik Asif & Amit Patel

Conference paper | First Online: 18 April 2019

1265 Accesses | 1 Citations

Part of the Lecture Notes in Mechanical Engineering book series (LNME)

Abstract

The aim of the study is to design and optimize crankshaft for a single-cylinder four-stroke overhead valve (OHV) spark ignition engine. This paper used reverse engineering techniques, in order to obtain an existing physical model. A three-dimensional crankshaft has been created with the help of SOLIDWORKS, and it is imported to ANSYS environment for the

Cite this paper

Pandiyan, A., Arun Kumar, G., Shaik Asif, S., Patel, A. (2019). Parametric Optimization of Single-Cylinder Four-Stroke Spark Ignition Engine Crankshaft Using Four Different Materials—AISI 1040, AISI 4340, AISI 4140 and AISI 4615. In: Hiremath, S., Shanmugam, N., Bapu, B. (eds) Advances in Manufacturing Technology. Lecture Notes in Mechanical Engineering. Springer, Singapore. https://doi.org/10.1007/978-981-13-6374-0_64

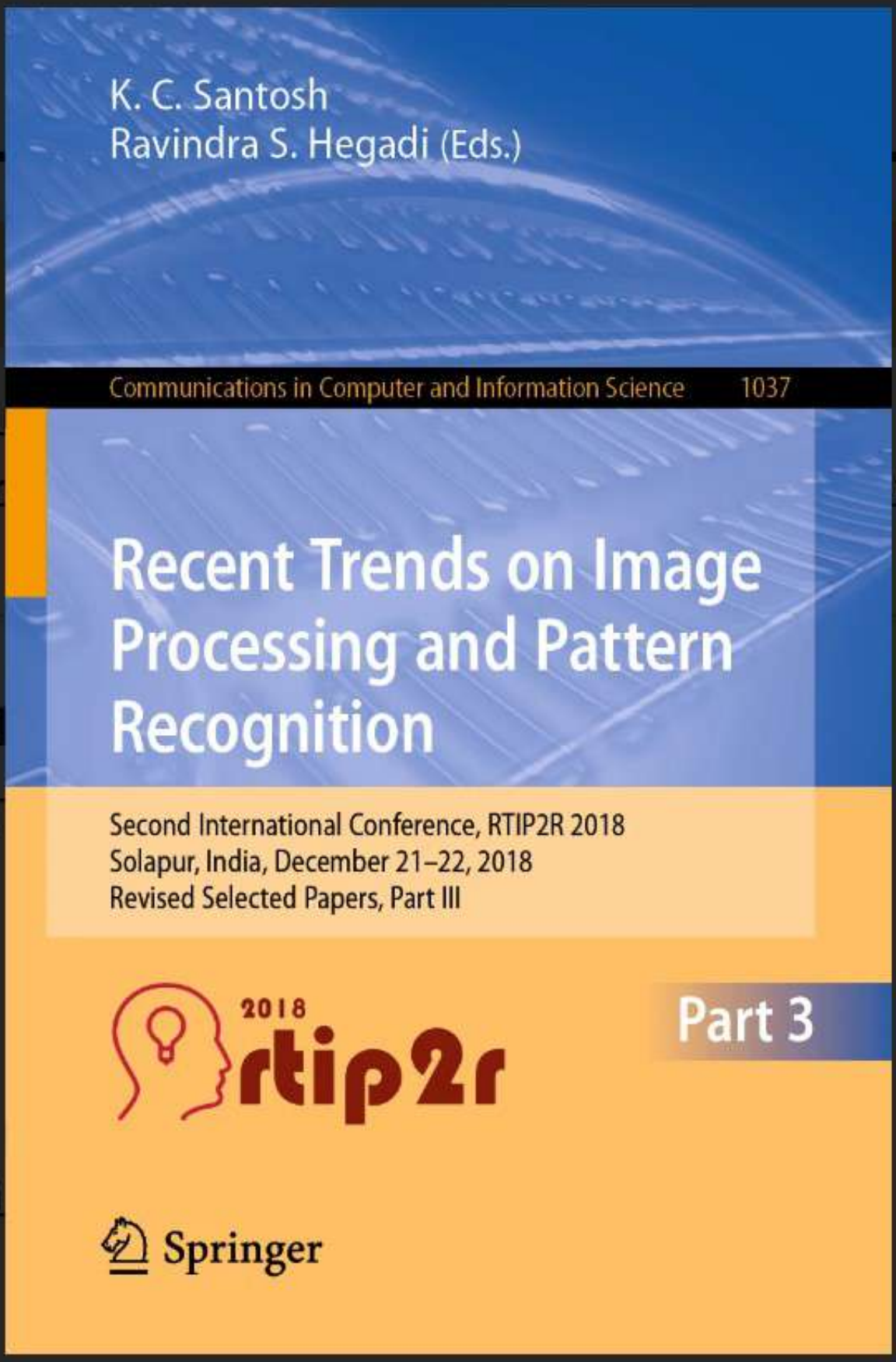
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-6374-0_64

Published	Publisher Name	Print ISBN
18 April 2019	Springer, Singapore	978-981-13-6373-3
Online ISBN	eBook Packages	
978-981-13-6374-0	Engineering Engineering (R0)	


K. C. Santosh
Ravindra S. Hegadi (Eds.)
Communications in Computer and Information Science 1037
Recent Trends on Image Processing and Pattern Recognition
Second International Conference, RTIP2R 2018
Solapur, India, December 21–22, 2018
Revised Selected Papers, Part III
2018 rtip2r
Part 3
Springer

On-Line Devanagari Handwritten Character Recognition Using Moments Features

Shalaka Prasad Deore & Albert Pravin

Conference paper | First Online: 17 July 2019

756 Accesses | **9** Citations

Part of the Communications in Computer and Information Science book series (CCIS,volume 1037)

Abstract

Now a days recognizing the handwritten character is receiving high significance because of numerous applications like Educational Software, On-line Signature Verification, Bank Cheque Processing, postal code recognition, Electronic library etc. Very less work is accounted in the research of Devanagari handwritten character recognition (HWDCR), so that there is a large scope of research in this area. In this paper we proposed a HWDCR system that recognizes Devanagari handwritten characters, the most popular script in India. Using pen tablet handwritten character is inputted and its on-line features are extracted like sequence of (x, y) coordinates, stroke and pressure information which are passed to classifier for classification.

Cite this paper

Deore, S.P., Pravin, A. (2019). On-Line Devanagari Handwritten Character Recognition Using Moments Features. In: Santosh, K., Hegadi, R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1037. Springer, Singapore. https://doi.org/10.1007/978-981-13-9187-3_4

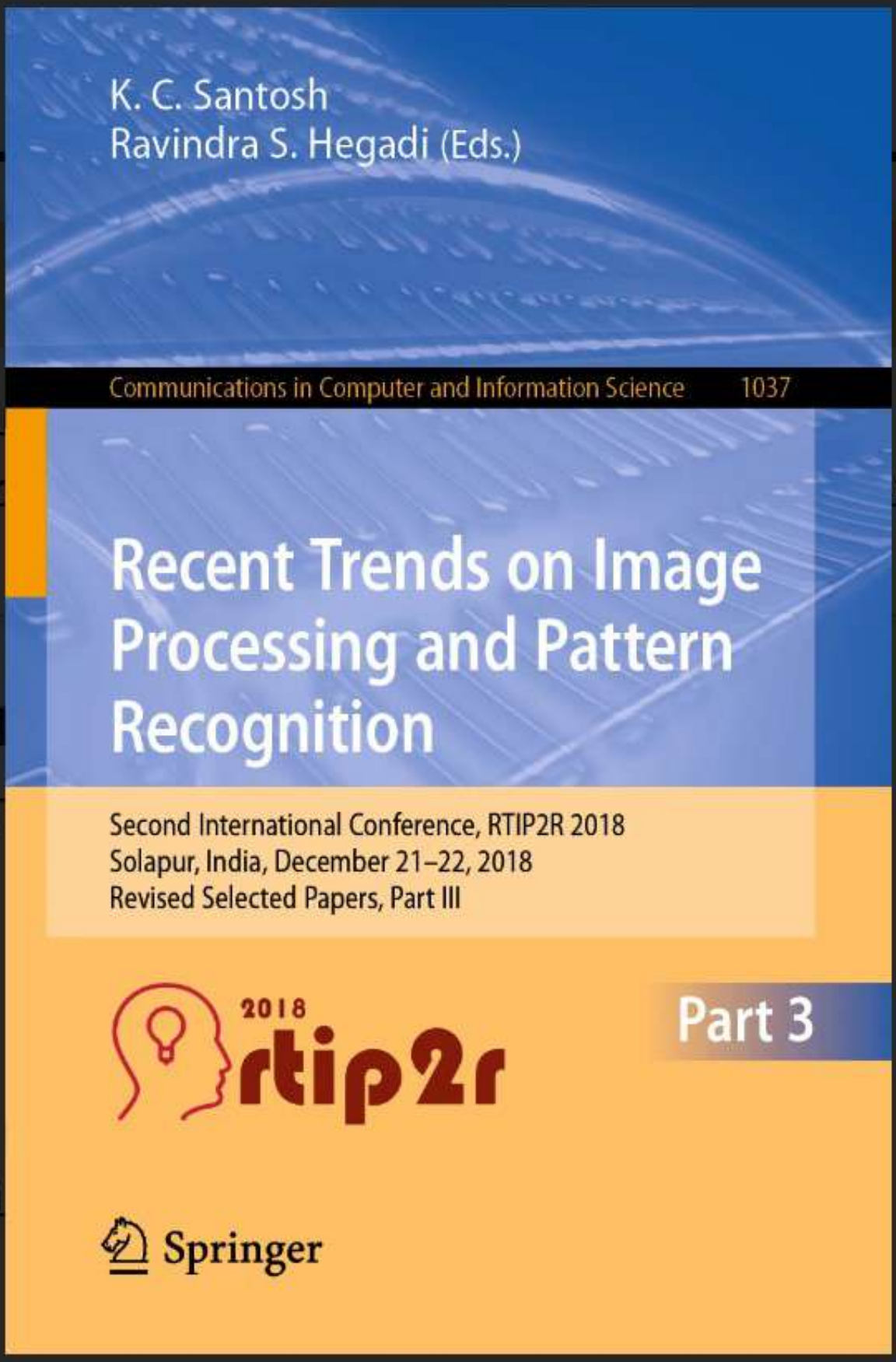
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-9187-3_4

Published	Publisher Name	Print ISBN
17 July 2019	Springer, Singapore	978-981-13-9186-6
Online ISBN	eBook Packages	
978-981-13-9187-3	Computer Science Computer Science (R0)	


K. C. Santosh
Ravindra S. Hegadi (Eds.)
Communications in Computer and Information Science 1037
Recent Trends on Image Processing and Pattern Recognition
Second International Conference, RTIP2R 2018
Solapur, India, December 21–22, 2018
Revised Selected Papers, Part III
2018 rtip2r
Part 3
Springer

On-Line Devanagari Handwritten Character Recognition Using Moments Features

Shalaka Prasad Deore & Albert Pravin

Conference paper | First Online: 17 July 2019

756 Accesses | **9** Citations

Part of the Communications in Computer and Information Science book series (CCIS,volume 1037)

Abstract

Now a days recognizing the handwritten character is receiving high significance because of numerous applications like Educational Software, On-line Signature Verification, Bank Cheque Processing, postal code recognition, Electronic library etc. Very less work is accounted in the research of Devanagari handwritten character recognition (HWDCR), so that there is a large scope of research in this area. In this paper we proposed a HWDCR system that recognizes Devanagari handwritten characters, the most popular script in India. Using pen tablet handwritten character is inputted and its on-line features are extracted like sequence of (x, y) coordinates, stroke and pressure information which are passed to classifier for classification.

Cite this paper

Deore, S.P., Pravin, A. (2019). On-Line Devanagari Handwritten Character Recognition Using Moments Features. In: Santosh, K., Hegadi, R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1037. Springer, Singapore. https://doi.org/10.1007/978-981-13-9187-3_4

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-9187-3_4

Published	Publisher Name	Print ISBN
17 July 2019	Springer, Singapore	978-981-13-9186-6
Online ISBN	eBook Packages	
978-981-13-9187-3	Computer Science Computer Science (R0)	

Advertisement



Hybrid Triodetection Approach: A Framework for Intrusion Detection

M. Mahithaa Sree, M. Saranya & S. Prayla Shyry

Conference paper | First Online: 21 December 2018

1360 Accesses | **2** Citations

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 26)

Abstract

With the emerging technologies in the internet, internal network security became a challenging issue in cyberspace. Though there are numerous devices and mechanisms to restrict the intruder attacks, the present inventions fall short of the mark due to the lack of accuracy. System access controls became a challenging dead weight even to the secure systems. The major productive defense today is use of intrusion detection system (IDS) and firewall. IDS works beyond the firewall by providing the protection to the network from

Cite this paper

Mahithaa Sree, M., Saranya, M., Prayla Shyry, S. (2019). Hybrid Triodetection Approach: A Framework for Intrusion Detection. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (eds) International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018. ICICI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-030-03146-6_119

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-03146-6_119

Published	Publisher Name	Print ISBN
21 December 2018	Springer, Cham	978-3-030-03145-9
Online ISBN	eBook Packages	
978-3-030-03146-6	Intelligent Technologies and	

Advertisement



Hybrid Triodetection Approach: A Framework for Intrusion Detection

M. Mahithaa Sree, M. Saranya & S. Prayla Shyry

Conference paper | First Online: 21 December 2018

1360 Accesses | **2** Citations

Part of the Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT,volume 26)

Abstract

With the emerging technologies in the internet, internal network security became a challenging issue in cyberspace. Though there are numerous devices and mechanisms to restrict the intruder attacks, the present inventions fall short of the mark due to the lack of accuracy. System access controls became a challenging dead weight even to the secure systems. The major productive defense today is use of intrusion detection system (IDS) and firewall. IDS works beyond the firewall by providing the protection to the network from

Cite this paper

Mahithaa Sree, M., Saranya, M., Prayla Shyry, S. (2019). Hybrid Triodetection Approach: A Framework for Intrusion Detection. In: Hemanth, J., Fernando, X., Lafata, P., Baig, Z. (eds) International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018. ICICI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 26. Springer, Cham. https://doi.org/10.1007/978-3-030-03146-6_119

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-3-030-03146-6_119

Published	Publisher Name	Print ISBN
21 December 2018	Springer, Cham	978-3-030-03145-9
Online ISBN	eBook Packages	
978-3-030-03146-6	Intelligent Technologies and	

Office Automation for Digital India

Publisher: IEEE

Cite This

PDF

Shashank Kumar ; Rohit Goyal ; T. Prem Jacob **All Authors**

101
Full
Text Views

Abstract

Document Sections

I. Introduction

Abstract:

Automation Plays very eminent role in our lives. It makes the work easier and more simpler. With the increase in consumption of energy and population there is humungous need of energy conservation in the workplace. One of the aspects of maintaining a productive workplace is making sure that there are effective health, safety and security procedures in

Shashank Kumar ; Rohit Goyal ; T. Prem Jacob **All Authors**

101
Full
Text Views

Abstract

Document Sections

I. Introduction

II. Proposed System

III. Architecture

IV. System Implementation

V. Material and Methodologies

Show Full Outline

Abstract:
Automation Plays very eminent role in our lives. It makes the work easier and more simpler. With the increase in consumption of energy and population there is humungous need of energy conservation in the workplace. One of the aspects of maintaining a productive workplace is making sure that there are effective health, safety and security procedures in place. After all there is a humungous amount of hard work and capitals are invested then why not on the security? And here this office Automation System plays an important role. This system is based on sub systems lighting, security and alarming system. The major reason of energy loss is lack in functionality and inability to access the appliances from remote locations. An android application is used to monitor and control the workplace from remote location.

Published in: 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR)

Published in: 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR)

Date of Conference: 11-13 July 2018

Date Added to IEEE *Xplore*: 11 November 2018

ISBN Information:

INSPEC Accession Number: 18233690

DOI: 10.1109/ICETIETR.2018.8529022

Publisher: IEEE

Conference Location: Ernakulam, India



International Conference on Computational Intelligence, Communications, and Business Analytics

↳ CICBA 2017: **Computational Intelligence, Communications, and Business Analytics** pp 206–216 | Cite as

Novel Methodology of Data Management in Ad Hoc Network Formulated Using Nanosensors for Detection of Industrial Pollutants

S. Gowri & J. Jabez

Conference paper | First Online: 24 September 2017

756 Accesses | **13** Citations

Part of the Communications in Computer and Information Science book series (CCIS,volume 775)

Access via your institution →

Sections Figures References

Abstract

References

Author information

Editor information

Rights and permissions

Activate Windows

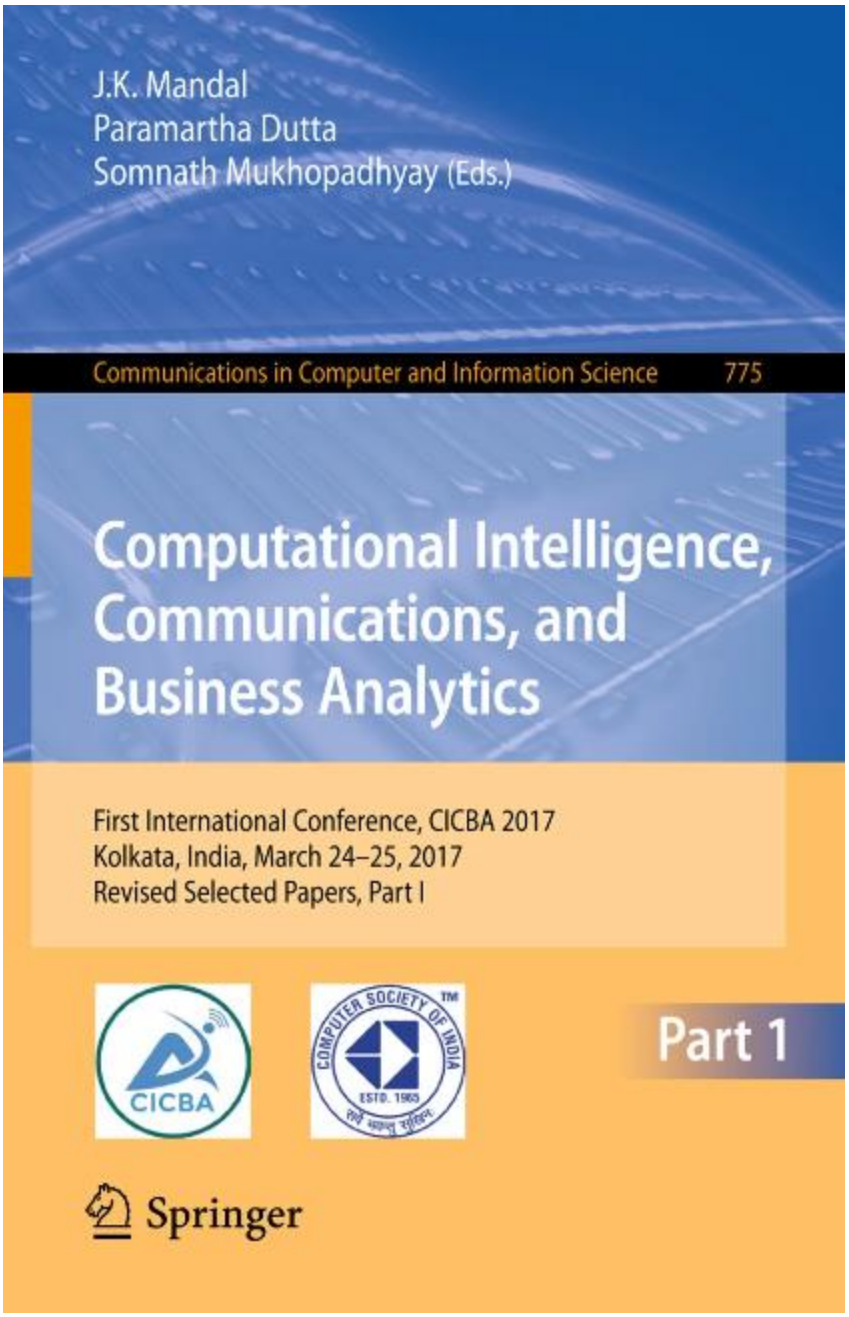
Abstract

In Ad hoc Network of Nanosensors for Wastage detection, clustering assist in nodal communication and in organization of the data fetched by the nanosensors in the network. The attempt of traditional cluster formation techniques degraded the formation of cluster in a precise manner. The data from the nanosensors which act as the nodes of the network have to be distinctively added into the clusters. The dynamic path selection cluster would achieve this distinct addition by dynamically creating a path to the data as an initial process and then redirecting the data to their appropriate cluster based to the readied scheme. The formation of clusters is evaluated on the bases of the time taken in formation of cluster and the relevance in information clustering. In this paper the dynamic path selection clustering algorithm is deployed to the cluster formation in data from sensor set for the detection of wastage from industries. Based on the data accumulated from the nanosensors the assignment of fetched information is clustered appropriately. The analyses of the resultant obtain on evaluation process shows the cluster formation life time and relevancy in cluster formation has been more feasible than that of the existing approaches.

Keywords

Ad hoc Network of Nanosensors **Industrial pollutants**

Dynamic path selection clustering **Data management**


J.K. Mandal
Paramartha Dutta
Somnath Mukhopadhyay (Eds.)
Communications in Computer and Information Science 775
Computational Intelligence, Communications, and Business Analytics
First International Conference, CICBA 2017
Kolkata, India, March 24–25, 2017
Revised Selected Papers, Part I
CICBA
COMPUTER SOCIETY OF INDIA
ESTD. 1965
Part 1
Springer

Notifying and filtering undesirable messages from online social network (OSN)

Publisher: IEEE

Cite This PDF

V. Vijeya Kaveri ; S. Gopal All Authors

4 Paper Citations
60 Full Text Views

V. Vijeya Kaveri ; S. Gopal All Authors

4 Paper Citations
60 Full Text Views

Abstract

Document Sections

I. Introduction

II. Related Work

III. Proposed System

IV. Mathematical Model

V. Results

Show Full Outline

Authors

Abstract:
Nowadays, we have a standard communicating source i.e. On-line Social Networks which are used to share, link and broadcast the substantial quantity of the data of human being's life. Constant and day-to-day communication infers the interchange of numerous kinds of information, comprising image, unrestricted text, video and audial information. A chief portion of social network content is established through the text of short, a notable sample are the mails eternally printed by users of OSN on specific private and public zones, named in common walls. This work recommends the system which imposes the screening process of content-based information and it is considered as a vital provision for OSN. This scheme permits the users of OSN to take a straight supervision on the messages which are mailed on the OSN user's walls. This is attained by the scheme of flexible rule-based, which permits a user to change the screening norms to be practiced to the OSN user's walls, and applied the tool of soft classifier i.e. Machine Learning spontaneously creating membership tags in provision of the process of filtering based on content.

Published in: International Confernce on Innovation Information in Computing Technologies

Published in: International Confernce on Innovation Information in Computing Technologies

Date of Conference: 19-20 February 2015

Date Added to IEEE *Xplore*: 04 February 2016

ISBN Information:

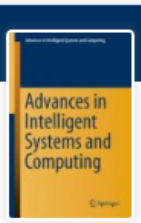
INSPEC Accession Number: 15774258

DOI: 10.1109/ICIICT.2015.7396091

Publisher: IEEE

Conference Location: Chennai, India

Springer



Book series

Advances in Intelligent Systems and Computing

Editorial board

About this book series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy

Abstract

The main objective of this paper is to design an information retrieval system for both text and image data using the concept of ontology. The focus of this paper is to improve information retrieval for sports events using Ontologies. For this purpose, an integration of domain knowledge with images using fuzzy ontology technique was implemented. In this work, the domain of basketball event is considered for creating low-level visual ontology for certain sport event images. The created multimodal ontology definition will provide a wide domain applicability, which allows the user to construct an ontology for any sport event. The domain ontologies are compared with the non-ontological information retrieval system to show the effectiveness of the ontological system.

Keywords

Ontology **Knowledge representation** **Semantic analysis** **OWL** **Machine learning** **Computer vision**

Springer Link

Search Log in



Proceedings of the International Conference on Soft Computing Systems pp 697–706 | Cite as

Multimodal Fuzzy Ontology Creation and Knowledge Information Retrieval

G. Nagarajan & R. I. Minu

Conference paper | First Online: 08 December 2015

1103 Accesses | **6** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 398)

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

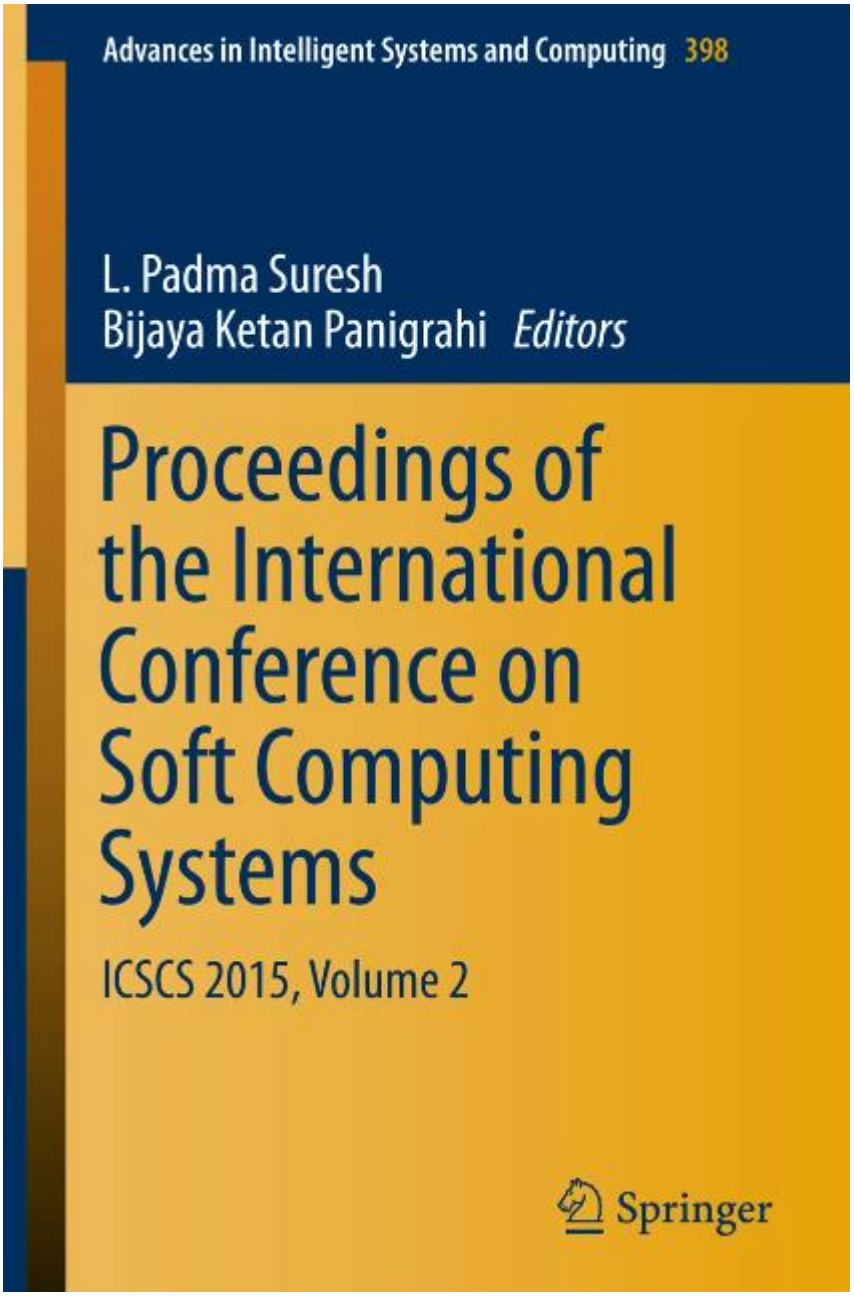
- DOI: 10.1007/978-81-322-2674-1_66
- Chapter length: 10 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

Activate Windows
Go to PC settings to activate Windows.

Abstract

The main objective of this paper is to design an information retrieval system for both text and


Advances in Intelligent Systems and Computing 398
L. Padma Suresh
Bijaya Ketan Panigrahi Editors
Proceedings of the International Conference on Soft Computing Systems
ICSCS 2015, Volume 2
Springer

Motion activated security camera using raspberry Pi

Publisher: IEEE

Cite This PDF

K. N. Karthick Kumar ; H. Natraj ; T. Prem Jacob **All Authors**

14 Paper Citations

1001 Full Text Views

Abstract

Document Sections

I. Introduction

Abstract:
Innovative headways have cleared path in private homes too. Shrewd homes are turning out to be all the more a need in this day and age. Control, observing and security are the vital elements of a savvy home and shrewd gadgets have a noteworthy impact in machine control. This paper focusses on the origination and execution of an inclusive continuous home security

Abstract:
Innovative headways have cleared path in private homes too. Shrewd homes are turning out to be all the more a need in this day and age. Control, observing and security are the vital elements of a savvy home and shrewd gadgets have a noteworthy impact in machine control. This paper focusses on the origination and execution of an inclusive continuous home security framework that is shrewd and totally deals with home security. This home security framework is named as Raspberry pi (RPI) and open CV with Harr cascade. This gadget can see any development and send warnings on the clients dashboard which thus advises the checking focus. Sensors are initiated amid your nonappearance or when you inform the arrangement of your nonattendance at home. It is conceivable to program the security framework to record exercise through the surveillance camera when a development is identified. The Raspberry Pi, a charge card measure pc with an inbuilt camera board can transform into a camera security framework when the camera board is put to utilize. This framework makes utilization of Open CV. The sensors identifying developments can respond to different situations. For example, movement in the family room, opening and shutting of entryways, breaking a window.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

Date Added to IEEE *Xplore*: 08 February 2018

ISBN Information:

INSPEC Accession Number: 17562499

DOI: 10.1109/ICCSP.2017.8286658

Publisher: IEEE

Conference Location: Chennai, India

Springer



Book series

Advances in Intelligent Systems and Computing

Editorial board

About this book series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 846)

Abstract

This paper is to design by adapting the Perturbation and Observation (P&O) algorithm of Maximum Power Point Tracking (MPPT) and intelligent controller to vary the duty ratio of KY boost converter using Photo Voltaic (PV) array as an energy input. This PV system extracts the energy from the sun, after the power converter module energy is fed to the battery for storage. The KY boost converter is used to boost the voltage to a certain level. The switches mainly work on MPPT conditions. This converter has higher voltage conversion ratio, low ripple, and continuous current. This combination makes the integration work perfectly under any transient disturbance condition occurs in the input side and it stays connected to the battery through charge controller. In the charge controller unit, PI controller is used to regulate the battery parameters. MPPT method-based Fuzzy Logic Controller (FLC) is implemented for PV panel by change in irradiance from (600 to 1000 W/M^2) with constant temperature at 25 °C.

Keywords

Fuzzy logic controller (FLC) **Perturbation and observation (P&O) algorithm** **Maximum Power Point Tracking (MPPT)** **PV (Photovoltaic) array** **KY boost converter**

Electronic ISSN
2194-5365

Print ISSN
2194-5357

Series Editor
Janusz Kacprzyk

Book titles in this series

Soft Computing Applications

Proceedings of the 9th International Workshop Soft Computing Applications (SOFA 2020)

Editors: Valentina Emilia Balas, Lakhmi C. Jain, Marius Mircea Balas & Dumitru Baleanu

Copyright: 2023

Soft cover | eBook



Springer Link



International Conference on Intelligent Computing and Applications pp 337–348 | Cite as

Maximum Power Tracking for PV Array System Using Fuzzy Logic Controller

M. Pushpavalli & N. M. Jothi Swaroopan

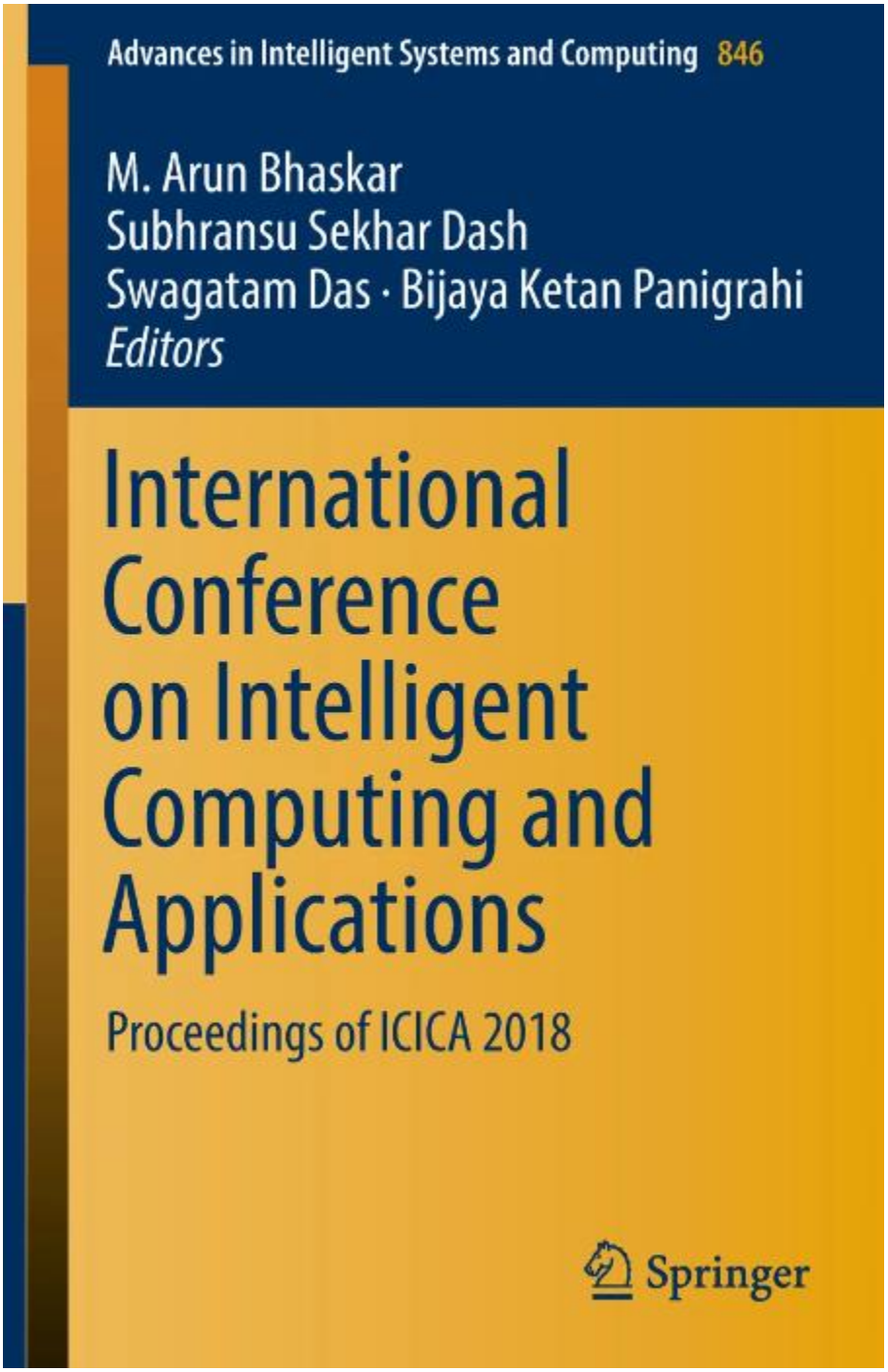
Conference paper | First Online: 09 September 2018

253 Accesses | **3** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 846)

Abstract

This paper is to design by adapting the Perturbation and Observation (P&O) algorithm of


Advances in Intelligent Systems and Computing 846
M. Arun Bhaskar
Subhransu Sekhar Dash
Swagatam Das · Bijaya Ketan Panigrahi
Editors
International Conference on Intelligent Computing and Applications
Proceedings of ICICA 2018
Springer

Low cost automated facial recognition system

Publisher: IEEE

Cite This PDF

M. Yashwanth Sai ; R Vijai Chandra Prasad ; P. R. Niveditha ; T. Sasipraba ; S. Vigneshwari ; S. Gowri **All Authors**

5 Paper Citations

583 Full Text Views

Abstract

Document Sections

I. Introduction

Abstract:
With the dawn of rising sophisticated technology where everything or rather everyone depends on smartphones, tablets and voice assistants. With such increase in the rate of people using technology to automate their tasks, it is impractical to hire a person for monitoring the CCTV or the IP camera feed for the intruders. Hence, the concept of

Abstract:
With the dawn of rising sophisticated technology where everything or rather everyone depends on smartphones, tablets and voice assistants. With such increase in the rate of people using technology to automate their tasks, it is impractical to hire a person for monitoring the CCTV or the IP camera feed for the intruders. Hence, the concept of automation can be implemented here as well. CCTV cameras present in a particular place are highly prevalent but then someone has to monitor the prospects of the recorded video. If an intruder comes in then, the person in charge has to raise the alarm to alert the security. Surveillance practices that are capable of facial recognition are becoming more common these days, but their prices remain high making it unaffordable for homes or small-scale businesses. So, the proposed facial recognition system is efficient as well as low-cost which not only recognizes the faces but also automates the task of alerting the user through e-mail and mobile notification so that no human intervention is required to alert the police. In addition to this, a log will also be generated with the details of the entry and the exit of the persons in the area where this system is installed. The proposed system uses a raspberry pi module and a camera module that fits in the CSI (camera serial interface) of the raspberry pi chip. Alternatively, a powerful USB camera can also be used to achieve the same task.

Published in: 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)

Published in: 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)

Date of Conference: 22-24 February 2017

Date Added to IEEE *Xplore*: 23 November 2017

ISBN Information:

INSPEC Accession Number: 17393497

DOI: 10.1109/ICECCT.2017.8117829

Publisher: IEEE

Conference Location: Coimbatore, India

Locker security system using facial recognition and One Time Password (OTP)

Publisher: IEEE Cite This PDF

N. Anusha ; A. Darshan Sai ; B. Srikar All Authors

N. Anusha
Computer Science and Engineering, Sathyabama University, Chennai, India

4 Paper Citations
529 Full Text Views

Abstract
Document Sections
I. Introduction
II. Motivation of the Research
III. Relatedk Work
IV. Proposed Method
V. Methodology
Show Full Outline
Authors

Abstract:
To overcome the security issues faced with the locker system now-a-days, the locker security system is proposed using IoT (Internet of Things), face recognition and OTP (One Time Password). Each time when an individual wants to access the locker, he/she should type the PIN (Personal Identification number) which is already set by them. If the PIN entered is valid, the camera connected to the locker captures and compares the image of the user using Eigen face and PCA (Principle Component Analysis) algorithms. Then the image will be matched with the database records by using Euclidean distance and Viola-Jones algorithm. Once the image gets matched One Time Password (OTP) will be generated and sent via SMS (Short Message Service) to the registered mobile using GSM (Global System for Mobile Communication) service and e-mail notification also will be sent. If the features are not matched then the log details related to the transaction will be sent to the registered number and to e-mail account. By using which the owner can easily complaint to the police officials about intruder by sending the log and location details via SMS.

Published in: 2017 International Conference on Wireless Communications, Signal Processing and Networking

Published in: 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)

Date of Conference: 22-24 March 2017

Date Added to IEEE *Xplore*: 22 February 2018

ISBN Information:

INSPEC Accession Number: 17598419

DOI: 10.1109/WiSPNET.2017.8299874

Publisher: IEEE

Conference Location: Chennai, India

All
Search within Publication
ADVANCED SEARCH
Browse Conferences > International Conference on Sc... > 2017 Third International Confe...
Copy Persistent Link | Browse Title List | Sign up for Conference Alerts
Proceedings | All Proceedings | Popular

International Conference on Science Technology Engineering and Management (ICONSTEM)

2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)
23-24 March 2017
DOI: 10.1109/ICONSTEM40713.2017

Search within results
Download PDFs | Per Page: 25 | Export | Email Selected Results

N. Anusha
Computer Science and Engineering, Sathyabama University, Chennai, India
uthors

4 Paper Citations
529 Full Text Views

Abstract
Document Sections
I. Introduction
II. Motivation of the Research
III. Relatedk Work
IV. Proposed Method
V. Methodology
Show Full Outline
Authors

Abstract:
To overcome the security issues faced with the locker system now-a-days, the locker security system is proposed using IoT (Internet of Things), face recognition and OTP (One Time Password). Each time when an individual wants to access the locker, he/she should type the PIN (Personal Identification number) which is already set by them. If the PIN entered is valid, the camera connected to the locker captures and compares the image of the user using Eigen face and PCA (Principle Component Analysis) algorithms. Then the image will be matched with the database records by using Euclidean distance and Viola-Jones algorithm. Once the image gets matched One Time Password (OTP) will be generated and sent via SMS (Short Message Service) to the registered mobile using GSM (Global System for Mobile Communication) service and e-mail notification also will be sent. If the features are not matched then the log details related to the transaction will be sent to the registered number and to e-mail account. By using which the owner can easily complaint to the police officials about intruder by sending the log and location details via SMS.

Published in: 2017 International Conference on Wireless Communications, Signal Processing and Networking

Select All on Page

Sort By: Relevance

Literature survey on existing analytical schemes to optimize the mining results using incremental map reduce

D. Hemalatha; B. Bharathi

2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)

Year: 2017 | Conference Paper | Publisher: IEEE

Cited by: Papers (2)

Abstract HTML

Literature Survey of Role Policy Access Using Cloud

Publisher: IEEE Cite This PDF

Santosh Krishna Naidu ; Khan Afzal Ayub ; S. Revathy All Authors

63
Full
Text Views

Abstract:
Cloud Computing is a technology that consists of an internal server and a common server which is used to store data in it in such a way that it can be accessed universally. Big data is a collection of a large data set or data modules. Various studies have been made where the successful integration of cloud and big data have been performed. The advantage of linking big data with cloud is that the data availability becomes global and could be accessed easily. This opens the data up for vandalism and could be easily manipulated. To ensure proper data integrity, different encryption techniques could be used to provide security to the data. Role-Based policy access is an access policy, which provides access to the end users based upon their attributes, which are used to supposedly authenticate the user. It gives the end user the privileges like file access, file modification and updating. This survey determines in various ways that role policy access using cloud has its own advantages and disadvantages.

Published in: 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)

Published in: 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)

Date of Conference: 11-12 May 2018

Date Added to IEEE *Xplore*: 02 December 2018

ISBN Information:

INSPEC Accession Number: 18290965

DOI: 10.1109/ICOEI.2018.8553790

Publisher: IEEE

Conference Location: Tirunelveli, India

Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage

Publisher: IEEE Cite This PDF

Cheng-Kang Chu ; Sherman S.M. Chow ; Wen-Guey Tzeng ; Jianying Zhou ; Robert H. Deng **All Authors**

134 Paper Citations

5950 Full Text Views

Abstract:

Data sharing is an important functionality in cloud storage. In this paper, we show how to securely, efficiently, and flexibly share data with others in cloud storage. We describe new public-key cryptosystems that produce constant-size ciphertexts such that efficient delegation of decryption rights for any set of ciphertexts are possible. The novelty is that one can aggregate any set of secret keys and make them as compact as a single key, but encompassing the power of all the keys being aggregated. In other words, the secret key holder can release a constant-size aggregate key for flexible choices of ciphertext set in cloud storage, but the other encrypted files outside the set remain confidential. This compact aggregate key can be conveniently sent to others or be stored in a smart card with very limited secure storage. We provide formal security analysis of our schemes in the standard model. We also describe other application of our schemes. In particular, our schemes give the first public-key patient-controlled encryption for flexible hierarchy, which was yet to be known.

Published in: IEEE Transactions on Parallel and Distributed Systems (Volume: 25 , Issue: 2, February 2014)

Published in: IEEE Transactions on Parallel and Distributed Systems (Volume: 25 , Issue: 2, February 2014)

Page(s): 468 - 477

INSPEC Accession Number: 13991688

Date of Publication: 11 April 2013

DOI: 10.1109/TPDS.2013.112

ISSN Information:

Publisher: IEEE

Series: Advances in Engineering Research

Proceedings of the International Conference for Phoenixes on Emerging Current Trends in Engineering and Management (PECTEAM 2018)

HOME
PREFACE
ARTICLES
AUTHORS
SESSIONS
ORGANIZERS
PUBLISHING INFORMATION

< PREVIOUS ARTICLE IN VOLUME
NEXT ARTICLE IN VOLUME >

Integration of Multi Access Control System with Door, Device & Location Automation & Remote Network Control

Authors
J S Vimali

Corresponding Author
J S Vimali

Activate Windows
Go to PC settings to activate Windows.

DOI
https://doi.org/10.2991/pecteam-18.2018.24 How to use a DOI?

Keywords
Android, ABE, Microcontroller, RMI, Bluetooth.

Abstract
Android is the most popular mobile operating system currently developed by Google, based on the Linux kernel. It is open source, thus the Operating System itself can be ported to different categories of devices like Watches (Wearables), Cars (Auto), TV, Gaming Consoles etc. The current devices are packed with all kinds of sensors and data collection abilities, users also store their personal information on the device. The project aims at bringing all the access controls at user's finger tips on his/her Android device.It can be used to switch on/off home appliances like lights, fans, etc. via the Bluetooth module (HC-05). The Application connects to a LAN Network of PC, and is able to Shutdown, Restart or Log off the PC in the LAN Network. In order to perform these remote tasks, the application makes use of Java RMI (Remote Method Invocation).Java RMI, is based on the Stub/Skeleton technology.

Copyright
© 2018, the Authors. Published by Atlantis Press.

Open Access
This is an open access article distributed under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Series: Advances in Engineering Research

Proceedings of the International Conference for Phoenixes on Emerging Current Trends in Engineering and Management (PECTEAM 2018)

< PREVIOUS ARTICLE IN VOLUME NEXT ARTICLE IN VOLUME >

Integration of Multi Access Control System with Door, Device & Location Automation & Remote Network Control

Authors
J S Vimali

Corresponding Author
J S Vimali

Available Online February 2018.

DOI
https://doi.org/10.2991/pecteam-18.2018.24 How to use a DOI?

Activate Windows
Go to PC settings to activate Wind



Conference proceedings | © 2019

Innovative Design, Analysis and Development Practices in Aerospace and Automotive Engineering (I-DAD 2018)

Volume 2

Editors: U. Chandrasekhar, Lung-Jieh Yang, S. Gowthaman

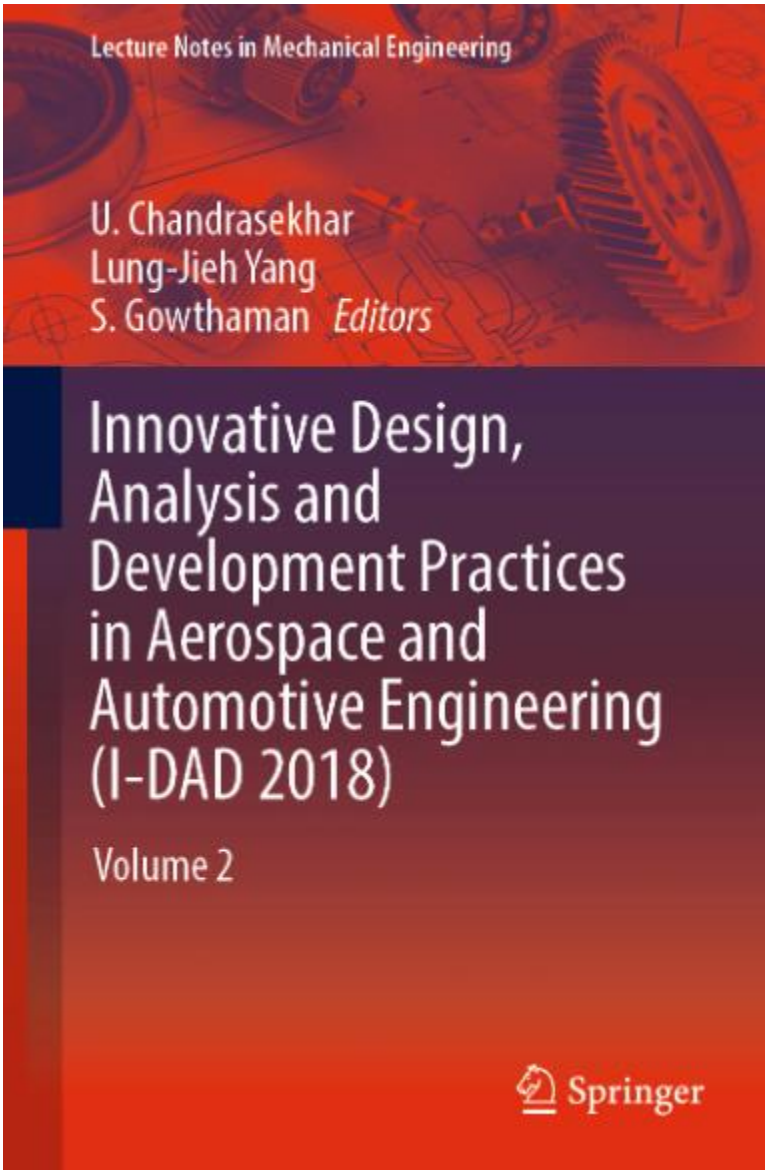
Provides a selection of the best papers from the I-DAD 2018 conference

Discusses new design concepts, analysis and manufacturing technologies for the aerospace and auto industry

Offers a valuable reference guide for research students and professionals

Part of the book series: Lecture Notes in Mechanical Engineering (LNME)

62k Accesses | **113** Citations


Lecture Notes in Mechanical Engineering
U. Chandrasekhar
Lung-Jieh Yang
S. Gowthaman Editors
Innovative Design, Analysis and Development Practices in Aerospace and Automotive Engineering (I-DAD 2018)
Volume 2
Springer

Improving the spectrum aware mobility management in cognitive radio network

Publisher: IEEE

Cite This PDF

K. Priyanka ; Kavitha Esther Rajakumari **All Authors**

174
Full
Text Views

Abstract

Document Sections

I Introduction

Abstract:
Wireless spread spectrum is a standout amongst the most valuable assets. With the fast development of the quantity of remote gadgets, interchanges over the unlicensed range groups get to be extremely swarmed. Data dissemination in the wireless environment is

Abstract:
Wireless spread spectrum is a standout amongst the most valuable assets. With the fast development of the quantity of remote gadgets, interchanges over the unlicensed range groups get to be extremely swarmed. Data dissemination in the wireless environment is usually characterized as the spreading of data to numerous goals through communicating. The principle target is to achieve the most extreme number of neighbors with each sent parcel. The main objective of the paper is to find the best routing and resource allocation strategies in order to minimize the average end-to-end delay of multiple data connections in the cognitive radio based wireless mesh network. Wireless mesh nodes utilize the cognitive radio to share the spectrum with primary users. Depending on the primary user's activities, the spectrum resources available to the cognitive mesh nodes are varying in both space and time. In this joint design scheme can accommodate the traffic load, or achieve half the delay compared to the disjoint methods.

Published in: 2017 International Conference on Information Communication and Embedded Systems (ICICES)

Published in: 2017 International Conference on Information Communication and Embedded Systems (ICICES)

Date of Conference: 23-24 February 2017 **INSPEC Accession Number:** 17261631

Date Added to IEEE *Xplore*: 19 October 2017 **DOI:** 10.1109/ICICES.2017.8070761

▸ **ISBN Information:**

Publisher: IEEE

Conference Location: Chennai, India

Springer Link

Search Log in



Advances in Smart Grid Technology pp 231–246 | Cite as

Implementing PV Energized SRSPM-Based Single-Phase Inverter for Induction Motor

P. Abirami, M. Pushpavalli & P. Sivagami

Conference paper | First Online: 23 September 2020

240 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 687)

Abstract

This research work explains in detail about the SRSPM technique along with solar energy. The

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-7245-6_19
- Chapter length: 16 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

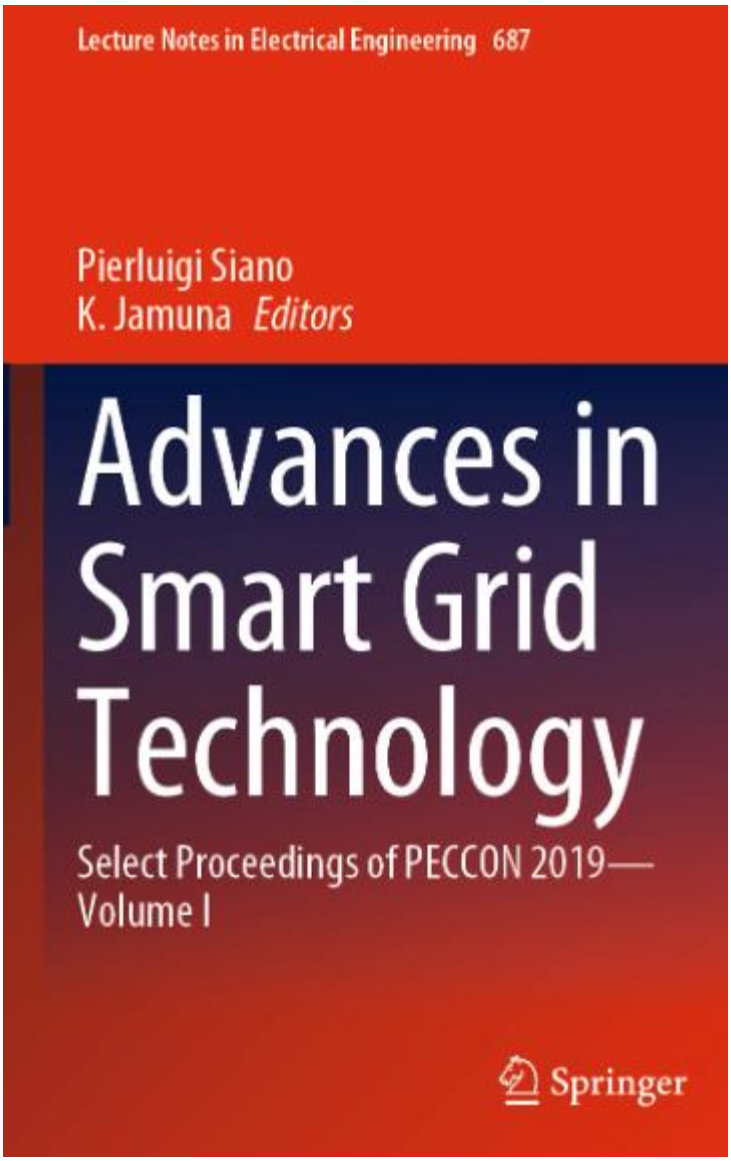
Activate Windows
Go to PC settings to activate Windows

Abstract

This research work explains in detail about the SRSPM technique along with solar energy. The drawback of the existing system is that it leads to high cost because of fuel cells and implements two isolated DC–DC converters in it which leads to high switching losses. Thus, to overcome the above said drawbacks, a single-reference six-pulse modulation technique along with solar power is implemented to improve the efficient usage of electrical vehicles. So, in this research work, solar energy is used to run the motor which is the renewable source energy. The PV panel generates DC supply, and a pulsating DC is produced by the hybrid technique SRSPM which is used to improve the efficiency of the vehicle. Here, the HF modulation is adopted to obtain compact, low cost, and lightweight system. Thus, the low voltage signal is amplified to the required three-phase voltage signals by means of a multistage inverter. HF inverter consists of a front-end DC–DC converter and three-phase PWM inverter. Here, two different modulations are implemented. SRSPM is used to control front-end full-bridge DC–DC converter to produce HF pulsating DC voltage having six-pulse information, and the average switching frequency and switching losses are reduced by balanced three-phase voltage produced by implementing 33% modulation technique of three-phase inverter. Thus, the proposed system simulation is carried out for three-phase supply and experimental setup is implemented for a single-phase induction motor.

Keywords

Fuel vehicles **Electrical vehicle** **SRSPM** **Induction motor first section**


Lecture Notes in Electrical Engineering 687
Pierluigi Siano
K. Jamuna Editors
Advances in Smart Grid Technology
Select Proceedings of PECCON 2019—Volume I
Springer

SEMANTIC SCHOLAR | Search 210,000,561 papers from all fields of science

DOI: 10.1109/ICCCSP.2018.8452861 • Corpus ID: 52161717

Implementation of One Time Password by 3*3 Vedic Multiplier

S. Shyry, M. Mahithaasree, M. Saranya • Published 1 February 2018 • Computer Science •
2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP)

Key Method

Although enormous methodologies are in market to generate One Time Password (OTP), day to day hackers are creating techniques to break the code and track the OTP. [...] Based on the credentials from the source and the length of the message, id of the client, two 8 bit binary values are created and converted to decimal through CB2D Procedure, to create two 3 digit decimal values. The decimal values are then multiplied by the 3×3 Vedic multiplier. The outcome of the Vedic multiplier is the One Time...

Expand

View on IEEE | doi.org | Save to Library | Create Alert | Cite

DOI: 10.1109/ICCCSP.2018.8452861 • Corpus ID: 52161717

Implementation of One Time Password by 3*3 Vedic Multiplier

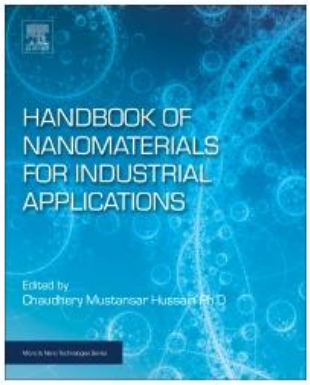
S. Shyry, M. Mahithaasree, M. Saranya • Published 1 February 2018 • Computer Science •
2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP)

Highlight Information | Methods | Results

Although enormous methodologies are in market to generate One Time Password (OTP), day to day hackers are creating techniques to break the code and track the OTP. In this paper, a coherent method of generating One Time Password is proposed. Based on the credentials from the source and the length of the message, id of the client, two 8 bit binary values are created and converted to decimal through CB2D Procedure, to create two 3 digit decimal values. The decimal values are then multiplied by the 3×3 Vedic multiplier. The outcome of the Vedic multiplier is the One Time Password. Also here, since the destination plays the role of showing the interest either to accept or deny the message from the source, the percentage of false authentication is reduced than the traditional techniques. Collapse

View on IEEE | doi.org | Save to Library | Create Alert | Cite

Book sale: save up to 30% on print and eBooks with free delivery. No promo code n



Preview

View on ScienceDirect

Handbook of Nanomaterials for Industrial Applications

1st Edition - July 19, 2018

☆☆☆☆☆ Write a review

Editor: Chaudhery Mustansar Hussain

Paperback ISBN: 9780128133514

eBook ISBN: 9780128133521

View series: Micro and Nano Technologies

Description

Handbook of Nanomaterials for Industrial Applications explores the use of

ScienceDirect

Journals &

View PDF | Access through your institution | Purchase PDF

Chapter contents | Book contents

Outline

Abstract

Keywords

52.1. Introduction

52.2. Engineered Nanomaterials

52.3. Impact of Nanomaterials on Industries

52.4. Safety Issues in Engineered Nanoparticles

52.5. Discussion and Conclusions

References

Show full outline

Cited By (1)

Handbook of Nanomaterials for Industrial Applications

Micro and Nano Technologies

2018, Pages 952-958



Chapter 52 - Impact of Engineered Nanomaterials for Environmental Industries

Vinita Vishwakarma

Show more

+ Add to Mendeley | Share | Cite

https://doi.org/10.1016/B978-0-12-813351-4.00054-7

Get rights and content

Abstract

The application of engineered nanomaterials (ENMs) is very broad with a multifaceted option for the industrial economy to revel. With such an immense potential, this book chapter ventures to present an overview on the impact of ENMs

+ Add to Mendeley Share Cite

https://doi.org/10.1016/B978-0-12-813351-4.00054-7

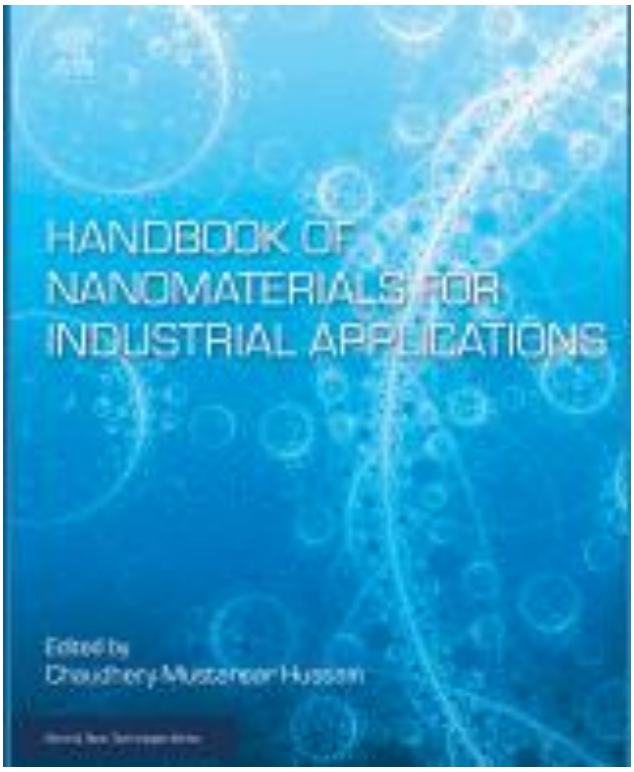
Get rights and content

Abstract

The application of engineered nanomaterials (ENMs) is very broad with a multifaceted option for the industrial economy to revel. With such an immense potential, this book chapter ventures to present an overview on the impact of ENMs on the wastewater treatment, solid waste, hazardous waste, leather waste, and other industrial process waste products. The synthesis, production, and disposal of engineered nanoparticles contaminate the environment. Many industries are unaware of the toxicity of nanomaterials and pay less attention to the negative aspect. Therefore, further studies are prompted to standardize the protocol for the manufacture, environmental release, and solution of nanomaterials. This book chapter addresses the impact and significant techniques of ENMs in the environmental industries, awareness of the impact on human health, handling emergencies, and eliminates its effects from the environment.

Previous chapter in book **Next** chapter in book

Keywords


HANDBOOK OF
NANOMATERIALS FOR
INDUSTRIAL APPLICATIONS
Edited by
Choudhery Mustansar Hussain

Identifying symptoms and treatment for heart disease from biomedical literature using text data mining

Publisher: IEEE

Cite This PDF

P. Sudeshna ; S. Bhanumathi ; M.R. Anish Hamlin **All Authors**

12 Paper Citations

423 Full Text Views

Abstract

Abstract:

Abstract:

As the heart diseases causing the major problem and they became the main cause of death worldwide, because it is very difficult for identifying the disease, based on symptoms, for that we need lots of experience and also knowledge. For knowing and identifying it is taking a lot of time for doctors also, because they have to observe the health condition and also the food habits of the patients, it is being a very long process, that the patient has to go to doctor and take the test and it is being a long process. The identification of the symptoms and also predicting the disease should be done fastly and efficiently, the problem of identifying and also predicting should be solved. we are proposing a technique, which predicts of the disease and probability of the disease and probability of the disease and probability of getting attacks, and there is predicted based on the patient data set. In our sector of health we have lots of information which is hidden and it may be useful for us in making the effective and also the precise decisions. Medicine Analysis is achieved by evidence using data mining Technique. Process is achieved by Analysis of Patient heart Health Condition, Formulating Questions gives heart diseases, Gathering evidences Analysis about heart disease. we modify an Automatic Machine Technique is used for Disease identification and correct Medicine Analysis based on evidence is achieved. Until Disease is not yet Diagnosis Evidence based Medicine Analysis is of no use. We also analyze the disease and best drug advised to that specific patient.

Published in: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Published in: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date of Conference: 22-23 March 2017

Date Added to IEEE *Xplore*: 15 February 2018

ISBN Information:

INSPEC Accession Number: 17580352

DOI: 10.1109/ICCPEIC.2017.8290359

Publisher: IEEE

Conference Location: Melmaruvathur, India

Identification of user POI in spatial data using Android implementations

Publisher: IEEE

Cite This PDF

S. Dhamodaran ; S. Manikandan ; P. Narendiran **All Authors**

77
Full
Text Views

Abstract

Document Sections

» Introduction

Abstract:
The disclosure of regions of passion for huge city areas is an crucial test. There is no mixture of Location centered search with POI is available. In particular, use novel query to implement the process, the path nearby cluster (PNC) query that identifies bundles of spatial content as

Abstract:
The disclosure of regions of passion for huge city areas is an crucial test. There is no mixture of Location centered search with POI is available. In particular, use novel query to implement the process, the path nearby cluster (PNC) query that identifies bundles of spatial content as for a customer indicated travel course. Spatial content can be purpose of top quality, e.g., stores, restaurants, and style stores, or they can be geo-labeled pictures, reduced in size range blog records, and twitter posts from place centered online social media, it is settings an Android operating system centered Application furthermore the consumer security is kept up for covering User place data.

Published in: 2016 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Published in: 2016 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date of Conference: 20-21 April 2016

Date Added to IEEE *Xplore*: 01 September 2016

ISBN Information:

INSPEC Accession Number: 16268122

DOI: 10.1109/ICCPEIC.2016.7557208

Publisher: IEEE

Conference Location: Melmaruvathur, India

Introduction

Hybrid scheduler to overcome the negative impact of job preemption for heterogeneous Hadoop systems

Publisher: IEEE

Cite This PDF

A. Uthirakumari ; P. Asha **All Authors**

5 Paper Citations

77 Full Text Views

Abstract

Document Sections

I. Introduction

Abstract:
To improve the resource utilization jobs running in the production clusters shares the available resources by assigning with different priorities. Preemption is one of the common technique used by the schedulers in Hadoop systems to prevent the delay in the job completion, also to

Abstract:
To improve the resource utilization jobs running in the production clusters shares the available resources by assigning with different priorities. Preemption is one of the common technique used by the schedulers in Hadoop systems to prevent the delay in the job completion, also to ensure that the available resource is effectively shared among the running jobs. In addition, this prevents large jobs, obtaining too many resources and making others to starve for resources. Jobs submitted to Hadoop systems may have a mixture of lengths, in such environment the existing preemption policies can delay the job completion resulting in the wasting of resources. The main aim of this paper is to size based scheduling using Hadoop Fair Sojourn Protocol with job preemption, which implements a size-based discipline that satisfies simultaneously system responsiveness and fairness requirements.

Published in: 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)

Published in: 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)

Date of Conference: 18-19 March 2016

INSPEC Accession Number: 16195172

Date Added to IEEE *Xplore*: 04 August 2016

DOI: 10.1109/ICCPCT.2016.7530206

ISBN Information:

Publisher: IEEE

Conference Location: Nagercoil, India

Home and office automation system using human activity recognition

Publisher: IEEE Cite This PDF

L. Mary Gladence ; Hari Haran Sivakumar ; Gobinath Venkatesan ; S. Shanmuga Priya All Authors

10 Paper Citations
228 Full Text Views

Conferences > 2017 International Conference...

Home and office automation system using human activity recognition

Publisher: IEEE Cite This PDF

L. Mary Gladence ; Hari Haran Sivakumar ; Gobinath Venkatesan ; S. Shanmuga Priya All Authors

10 Paper Citations
228 Full Text Views

Abstract

Document Sections

I. Introduction

II. Related Work

III. Proposed Work

IV. System Architecture

V. Implementation and

Abstract:
The recent years have witnessed significant progress in the automation of human activity recognition in order to realize intelligent environments which are capable of detecting users actions and gestures so that the needed services can be provided automatically and instantly for maximizing the user comfort and safety as well as minimizing energy. If a person enters into the room and sits/stands/lies down then the fan or AC should be automatically turned on. If a person comes and walks away or keeps roaming around, then they should not be turned on thus saving energy.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

INSPEC Accession Number: 17562427

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286463

ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India

High-resolution flood hazard mapping using remote sensing data

Publisher: IEEE

Cite This

PDF

S. Dhamodaran ; A. Shrthi ; Adline Suresh Thomas **All Authors**

1 Paper Citation

111 Full Text Views

1 Paper Citation

111 Full Text Views

Abstract

Document Sections

I. Introduction

II. Associated Work

III. Organization Model

IV. Performance Evaluation

V. Conclusion

Authors

Figures

References

Abstract:

The ability to look at An overflow of a large amount of water beyond its normal limitss with sensors hooked up on craft and satellites has been illustrious for a couple of years. Early launches of satellites and therefore the supply of aerial photography allowed investigation of the competencies to help An overflow of a large amount of water beyond its normal limits observance fromas manner as area. There may be right now aagreement among area organizations to streng then the aid that satellites will provide. This research is very main growth has been fostering our figuring out of the that for the period of which inside which faraway sensing can aid and even improve An overflow of a large amount of water beyond its normal limits modeling. This evaluation goes substantially further than utilizing a wet/dry An overflow of a large amount of water beyond its normal limits map for mannequin validation as in early reports of this type. A replacement methodology for finding An overflow of a large amount of water beyond its normal limitsendangerment with aid of remote sensing observations and hydraulics modeling used to be provided. The notion in the back of this strategy is that hydraulic variables appreciate from continuous spatially-distributed fluid mechanics modeling and and determined as separate satellite tv for pc-derived An overflow of a large amount of water beyond its normal limits extents ar correlate in time, so as that probabilities are transferred from the mannequin series to the observations.

Published in: 2016 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date of Conference: 20-21 April 2016

Date Added to IEEE *Xplore*: 01 September 2016

ISBN Information:

INSPEC Accession Number: 16268126

DOI: 10.1109/ICCPEIC.2016.7557210

Publisher: IEEE

Conference Location: Melmaruvathur, India

Springer Link



International Conference on Soft Computing Systems
↳ ICSCS 2018: **Soft Computing Systems** pp 200–210 | Cite as

Heart Block Recognition Using Image Processing and Back Propagation Neural Networks

P. Asha ✉, B. Sravani & P. SatyaPriya

Conference paper | First Online: 25 September 2018

1303 Accesses

Part of the Communications in Computer and Information Science book series (CCIS,volume 837)

Heart Block Recognition Using Image Processing and Back Propagation Neural Networks

P. Asha ✉, B. Sravani & P. SatyaPriya

Conference paper | First Online: 25 September 2018

1303 Accesses

Part of the Communications in Computer and Information Science book series (CCIS,volume 837)

Abstract

Nowadays in the health economy scenario, there has been a significant number of medicinal services data consisting of concealed details. The treatment for ischemic heart disease is a tough nut to crack as it calls for a widened comprehension and skill. One of the widely used conventional approaches to figure out heart ailment is through a doctor's assessment. Rather the disease can also be found out through a range of health checkups that includes Heart Magnetic Resonance Imaging (MRI) test, to arrive at an enhanced and useful output, original techniques are recommended, and the suggested methods are DTCWT and neural network (NN). The aforementioned method paves the way for an enhanced yield given identifying heart disease. A standout amongst the most regularly utilized image preparing system is Discrete Wavelet Transform (DTCWT) which best fits for changing images from the spatial domain into the frequency domain. The NN prepares the image with the help of extorted facet DTCWT feature. In this paper, the literature survey and the proposed methodologies, advantages and disadvantages of the proposed methodologies has been discussed. This is implemented in matlab.

Springer Link

Search Log in



International Conference on Soft Computing Systems

↳ ICSCS 2018: **Soft Computing Systems** pp 200–210 | Cite as

Heart Block Recognition Using Image Processing and Back Propagation Neural Networks

P. Asha, B. Sravani & P. SatyaPriya

Conference paper | First Online: 25 September 2018

1303 Accesses

Part of the Communications in Computer and Information Science book series (CCIS,volume 837)

Abstract

Nowadays in the health economy scenario, there has been a significant number of medicinal services data consisting of concealed details. The treatment for ischemic heart disease is a tough nut to crack as it calls for a widened comprehension and skill. One of the widely used conventional approaches to figure out heart ailment is through a doctor's assessment. Rather

Access via your institution →

Chapter EUR 29.95

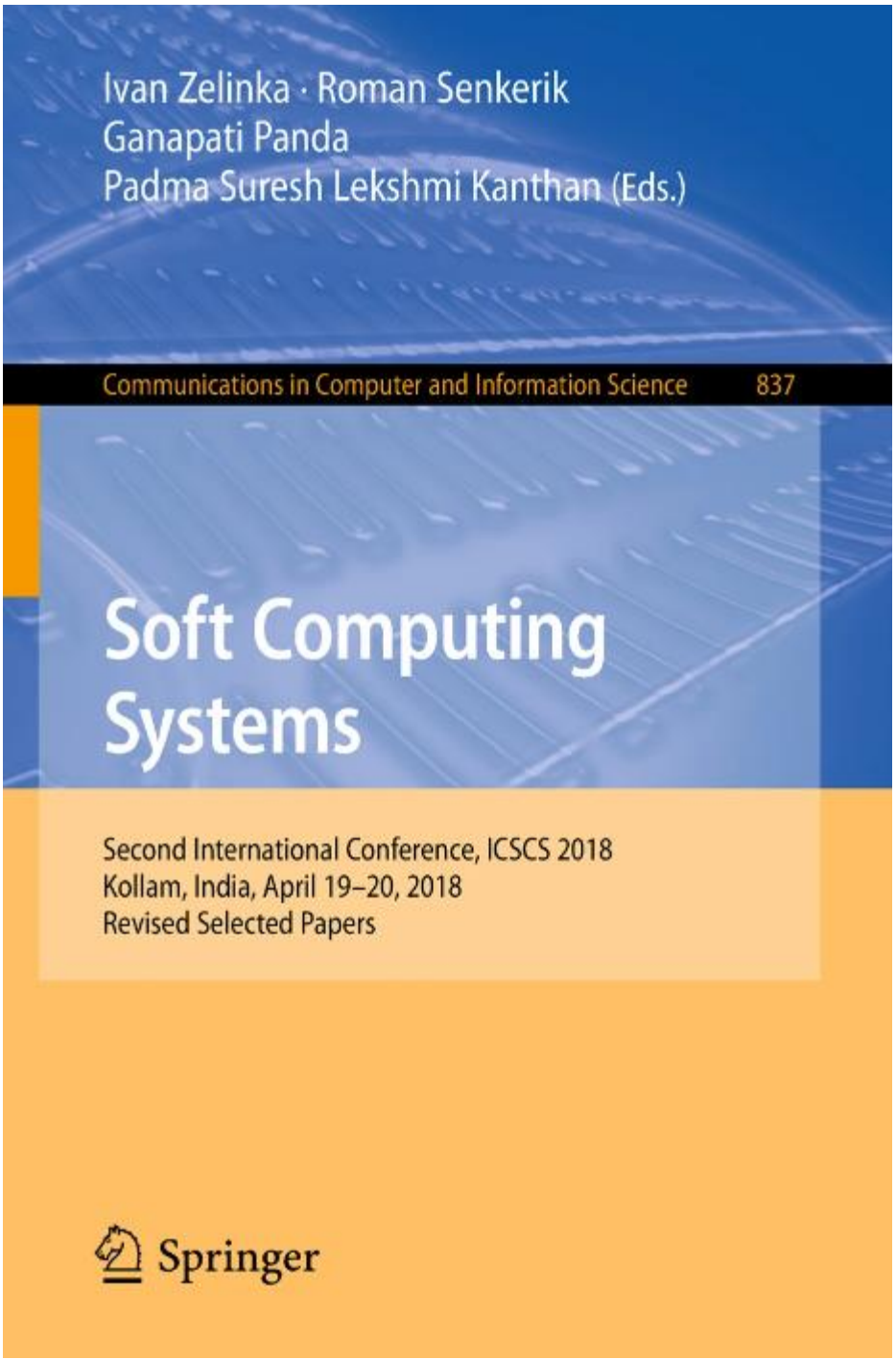
Price includes VAT (India)

- DOI: 10.1007/978-981-13-1936-5_23
- Chapter length: 11 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 96.29

Softcover Book EUR 115.99


Ivan Zelinka · Roman Senkerik
Ganapati Panda
Padma Suresh Lekshmi Kanthan (Eds.)
Communications in Computer and Information Science 837
Soft Computing Systems
Second International Conference, ICSCS 2018
Kollam, India, April 19–20, 2018
Revised Selected Papers
Springer

Health analysis of bicycle rider and security of bicycle using IoT

Publisher: IEEE Cite This PDF

Saumya Nath ; Sneha Sinha ; L. Mary Gladence ; Y. BevishJinila ; V. Rajalakshmi All Authors

8 Paper Citations
540 Full Text Views

Abstract
Document Sections
I. Introduction
II. Modules Used
III. Components Used
IV. Life Cycle
V. Result and Analysis

Abstract:
Internet of Things (IoT) is a new and growing field and has been used in various aspects including safe home, smart and hospitals. The same can be used for fitness tracking as well. This paper proposed a noveldevice for the health analysis of bicycle rider and security of bicycle. The device will help in keeping track of fitness of bicycle rider and make it hassle free. It will measure the heart rate of the rider, distance travelled and the speed of the bicycle to estimate the amount of energy consumed during the ride. The safety of the bicycles is also a great deal. The device uses antitheft technologies like keyless authentication using Bluetooth and accelerometer to detect theft attempt while the bicycle is in thelock position. The system relies on different, yet complimentary technologies like photoplethysmography, Reed switch, accelerometer, stepper motor, all interfaced using serial communication. The system connects to a smartphone with an app using Bluetooth. The mobile application stores and displays the data from the device in real time, perform analytics & provides security.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

INSPEC Accession Number: 17562548

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286475

ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India

Health analysis of bicycle rider and security of bicycle using IoT

Publisher: IEEE Cite This PDF

Saumya Nath ; Sneha Sinha ; L. Mary Gladence ; Y. BevishJinila ; V. Rajalakshmi **All Authors**

8 Paper Citations

540 Full Text Views

Conferences > 2017 International Conference...

Health analysis of bicycle rider and security of bicycle using IoT

Publisher: IEEE Cite This PDF

Saumya Nath ; Sneha Sinha ; L. Mary Gladence ; Y. BevishJinila ; V. Rajalakshmi All Authors

8 Paper Citations

540 Full Text Views

Abstract

Document Sections

I. Introduction

II. Modules Used

III. Components Used

IV. Life Cycle

V. Result and Analysis

Abstract:

Internet of Things (IoT) is a new and growing field and has been used in various aspects including safe home, smart and hospitals. The same can be used for fitness tracking as well. This paper proposed a noveldevice for the health analysis of bicycle rider and security of bicycle. The device will help in keeping track of fitness of bicycle rider and make it hassle free. It will measure the heart rate of the rider, distance travelled and the speed of the bicycle to estimate the amount of energy consumed during the ride. The safety of the bicycles is also a great deal. The device uses antitheft technologies like keyless authentication using Bluetooth and accelerometer to detect theft attempt while the bicycle is in thelock position. The system relies on different, yet complimentary technologies like photoplethysmography, Reed switch, accelerometer, stepper motor, all interfaced using serial communication. The system connects to a smartphone with an app using Bluetooth. The mobile application stores and displays the data from the device in real time, perform analytics & provides security.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

INSPEC Accession Number: 17562548

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286475

ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India

Springer Link

Search Log in



Microelectronics, Electromagnetics and Telecommunications pp 223–234 | Cite as

GSM Entrenched Intelligent Phase Changer

P. Sivagami & R. Harikrishnan

Conference paper | First Online: 03 November 2018

771 Accesses

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 521)

Abstract

Generally, the power fluctuations in the distribution system and its percentage of occurrence will be higher in one or two phases compared to all the three phases. Thus, in any commercial or domestic power supply system where three phases are available, it is advisable to have an

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-13-1906-8_24
- Chapter length: 12 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

Activate Windo
Go to Settings to act

Abstract

Generally, the power fluctuations in the distribution system and its percentage of occurrence will be higher in one or two phases compared to all the three phases. Thus, in any commercial or domestic power supply system where three phases are available, it is advisable to have an automatic changeover to the healthier phase for the single-phase critical loads. A microcontroller-based intelligent circuit is designed by using embedded system technology, to monitor and compare the voltage, frequency, and current with the predefined reference values. A single-phase critical load will be operational in any one of the phases. If the parameters monitored fall below or above the reference value in the loaded phase, then the load will automatically switch over to the next available healthier phase. A fault message will be sent to the intended recipients through SMS using GSM technology. An LCD display will show all the values including healthier also. Provision for an inverter backup is provided to give continuity of power to the critical load if in case all the three phases have fluctuations or power failures. Options are provided to control the phase changer by an authorized person through SMS also. Further, the under or overvoltage can be automatically controlled to a stable reference voltage level by means of SMS.

Keywords

GSM **PIC microcontroller** **Phase changer** **Three phase** **Control switch**

Lecture Notes in Electrical Engineering 521

Ganapati Panda
Suresh Chandra Satapathy
Birendra Biswal · Ramesh Bansal
Editors

Microelectronics, Electromagnetics and Telecommunications

Proceedings of the Fourth ICMEET 2018

Springer

Springer Link

Search Log in



International Conference on Intelligent Computing and Applications pp 371–379 | Cite as

Fuzzy Logic Controller-Based Bridgeless Cuk Converter for Power Factor Correction

R. Meena Devi, M. Kavitha, V. Geetha & M. Preethi Pauline Mary

Conference paper | First Online: 09 September 2018

281 Accesses | 1 Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 846)

Access via your institution

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-13-2182-5_35
- Chapter length: 9 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

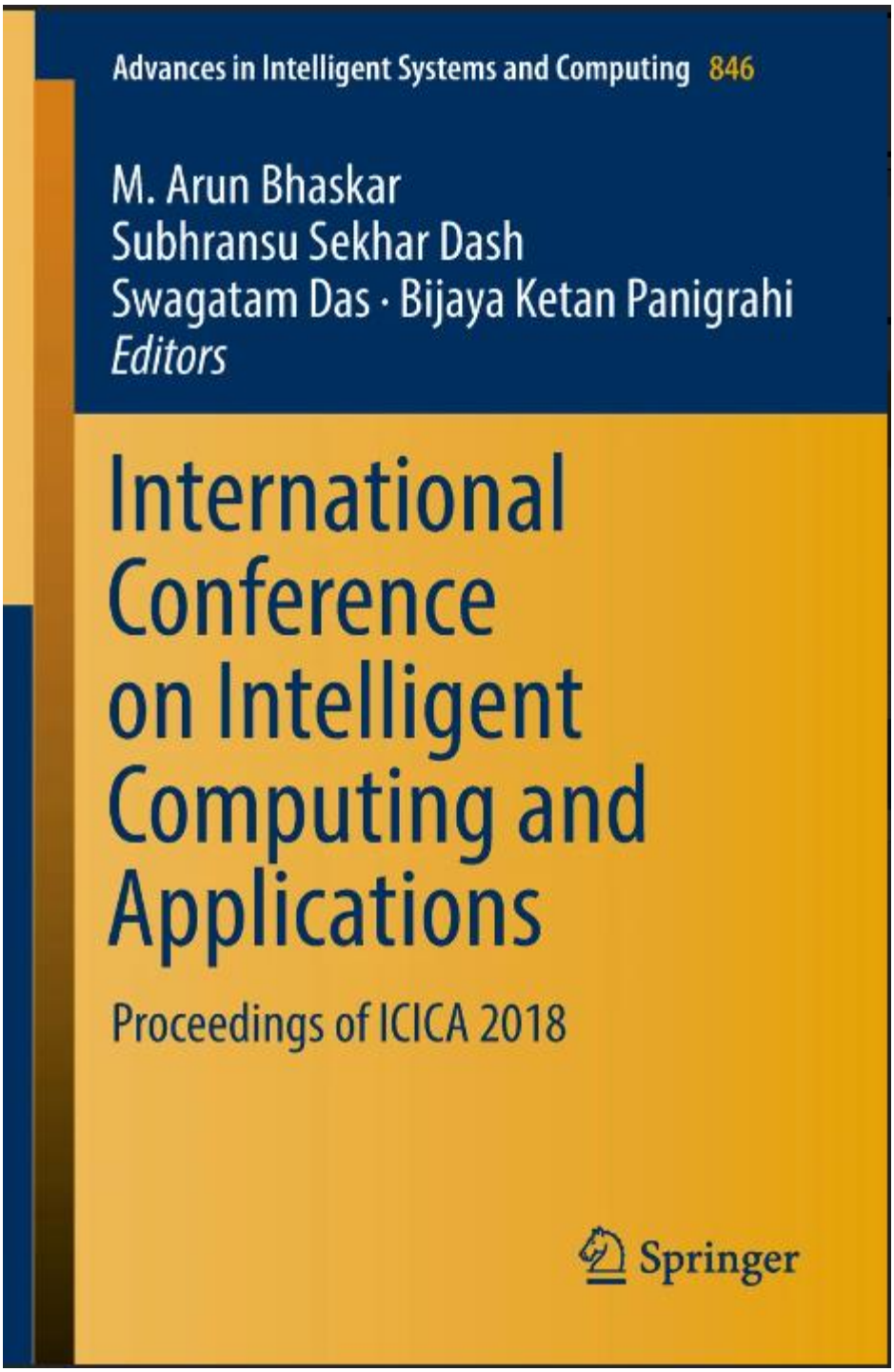
eBook EUR 139.09

Abstract

The Fuzzy logic controller-based single-phase bridgeless Cuk PFC converter with positive output voltage is proposed in this paper. This converter is used for both buck and boost operations. In case of low voltage output applications, it acts in buck mode which converts high input voltage to low output voltage. Whereas for applications involving high output voltage, it operates in boost mode which converts low input voltage into high output voltage. The conduction losses are reduced and efficiency of this system can be improved due to the absence of bridge diodes. Since bridge diodes are absent, the rectifier has a single switch, thus reducing the complexity of the circuit. The current loop circuit is not required, as discontinuous mode of operation is involved in rectifier. The simulation is put into execution by using MATLAB/Simulink software. The results of both open and closed loop were executed and compared. In closed loop, Fuzzy Logic Controller (FLC) provides reduced output voltage ripple with less settling time and better power factor.

Keywords

Cuk | Bridgeless | PFC | FLC | Positive output voltage


Advances in Intelligent Systems and Computing 846
M. Arun Bhaskar
Subhransu Sekhar Dash
Swagatam Das · Bijaya Ketan Panigrahi
Editors
International Conference on Intelligent Computing and Applications
Proceedings of ICICA 2018
Springer

Springer Link

Search Log in



Proceedings of the International Conference on Soft Computing Systems pp 167–177 | Cite as

Fuzzy-Based Multiloop Interleaved PFC Converter with High-Voltage Conversion System

P. Elangomenan & G. Nagarajan

Conference paper | First Online: 29 December 2015

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 397)

Abstract

This paper presents the implementation of an interleaved boost converter controlled by fuzzy logic control, combined with conventional multiloop control and interleaved pulse-width modulation. The conventional method has the drawback of increased current and voltage ripple. The generated modified proposed system will reduce the voltage and current ripple

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-81-322-2671-0_16
- Chapter length: 11 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

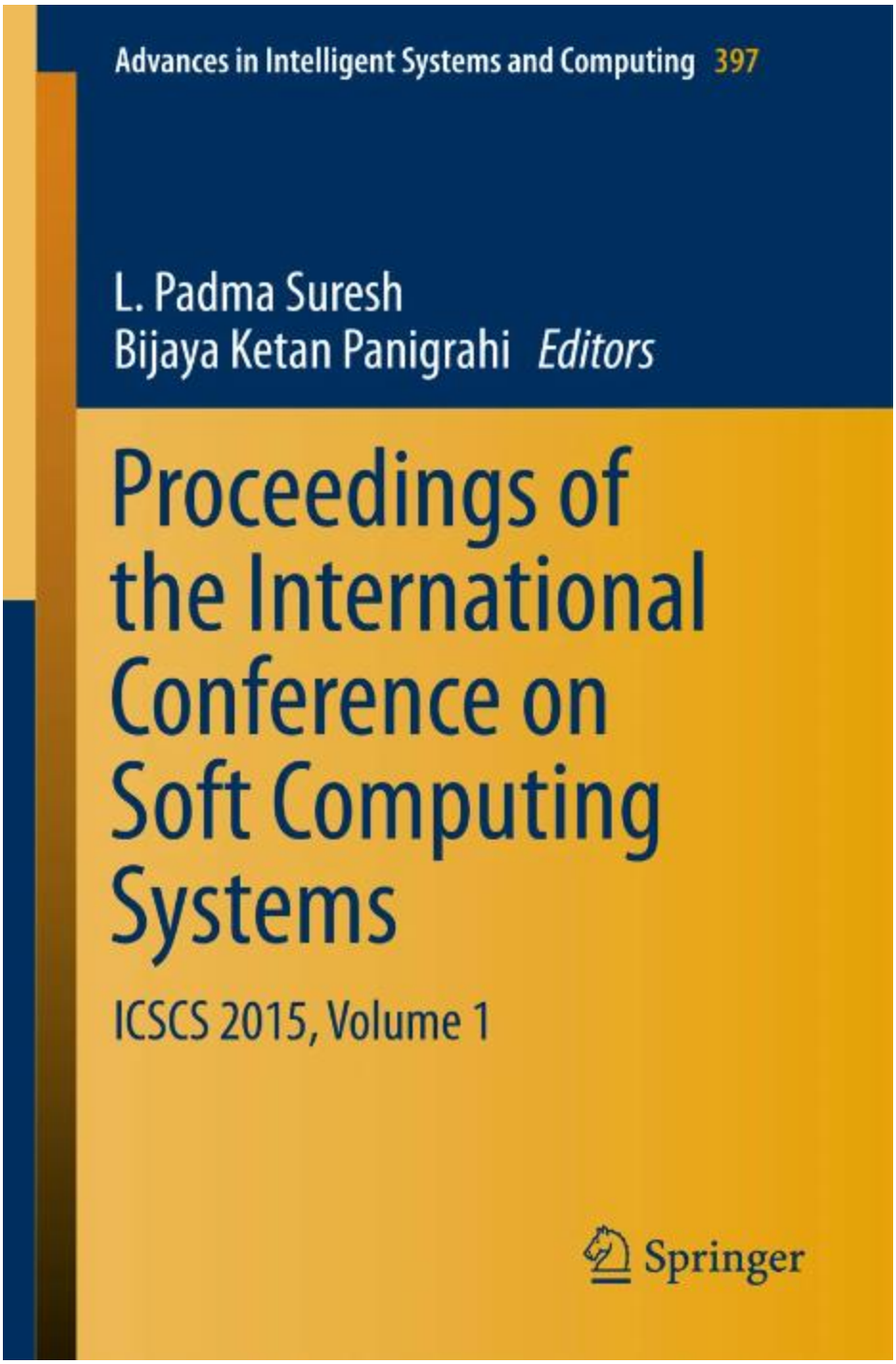
eBook EUR 245.03

Abstract

This paper presents the implementation of an interleaved boost converter controlled by fuzzy logic control, combined with conventional multiloop control and interleaved pulse-width modulation. The conventional method has the drawback of increased current and voltage ripple. The generated modified proposed system will reduce the voltage and current ripple and improve PFC correction. The average behavior of the interleaved three-level switch-mode rectifier (SMR) is similar to the voltage stress and can be reduced using ZVS and ZCS soft switching method. The proposed method is simulated using MATLAB software with embedded controller in the hardware; both the results are verified by this system.

Keywords

Fuzzy logic **Interleaved boost converter (ILBC)** **PI controller** **PWM technique** **Switch-mode rectifier (SMR)**


Advances in Intelligent Systems and Computing 397
L. Padma Suresh
Bijaya Ketan Panigrahi Editors
Proceedings of the International Conference on Soft Computing Systems
ICSCS 2015, Volume 1
Springer

Springer Link

Search Log in



Proceedings of the International Conference on Soft Computing Systems pp 235–247 | Cite as

FSM-Based VLSI Architecture for the 3 × 3 Window-Based DBUTMPF Algorithm

K. Vasanth, V. Elanangai, S. Saravanan & G. Nagarajan

Conference paper | First Online: 08 December 2015

1084 Accesses | **7** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 398)

Abstract

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-81-322-2674-1_24
- Chapter length: 13 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Abstract

This paper gives a Novel FSM-based architecture for the decision-based unsymmetrical trimmed midpoint algorithm used for high-density salt and pepper noise (SPN) removal in images. The proposed VLSI architecture uses a FSM-based scheduler for the evaluation of unsymmetrical trimmed midpoint in a fixed 3 × 3 window. The proposed scheduler moves between 4 states of the finite state machine for the evaluation of a suitable value to replace the corrupted value in the decision-based algorithm. The proposed architecture consists of sorting network, FSM scheduler and decision unit, which uses 9 values of the current processing window. This setup acts as a sequential architecture. During the simulation the first output of the decision appears after 16 clock cycles. The proposed architecture was targeted for Xc3e5000-5fg900 FPGA and the proposed architecture occupies 857 slices, consumes 298 MW power, and operates at 98.38 MHz frequency.

Keywords

FSM-based architecture **Decision-based unsymmetrical trimmed midpoint algorithm** **Sequential architecture**

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-81-322-2674-1_24
- Chapter length: 13 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 192.59

Softcover Book EUR 229.99

Learn about institutional subscriptions

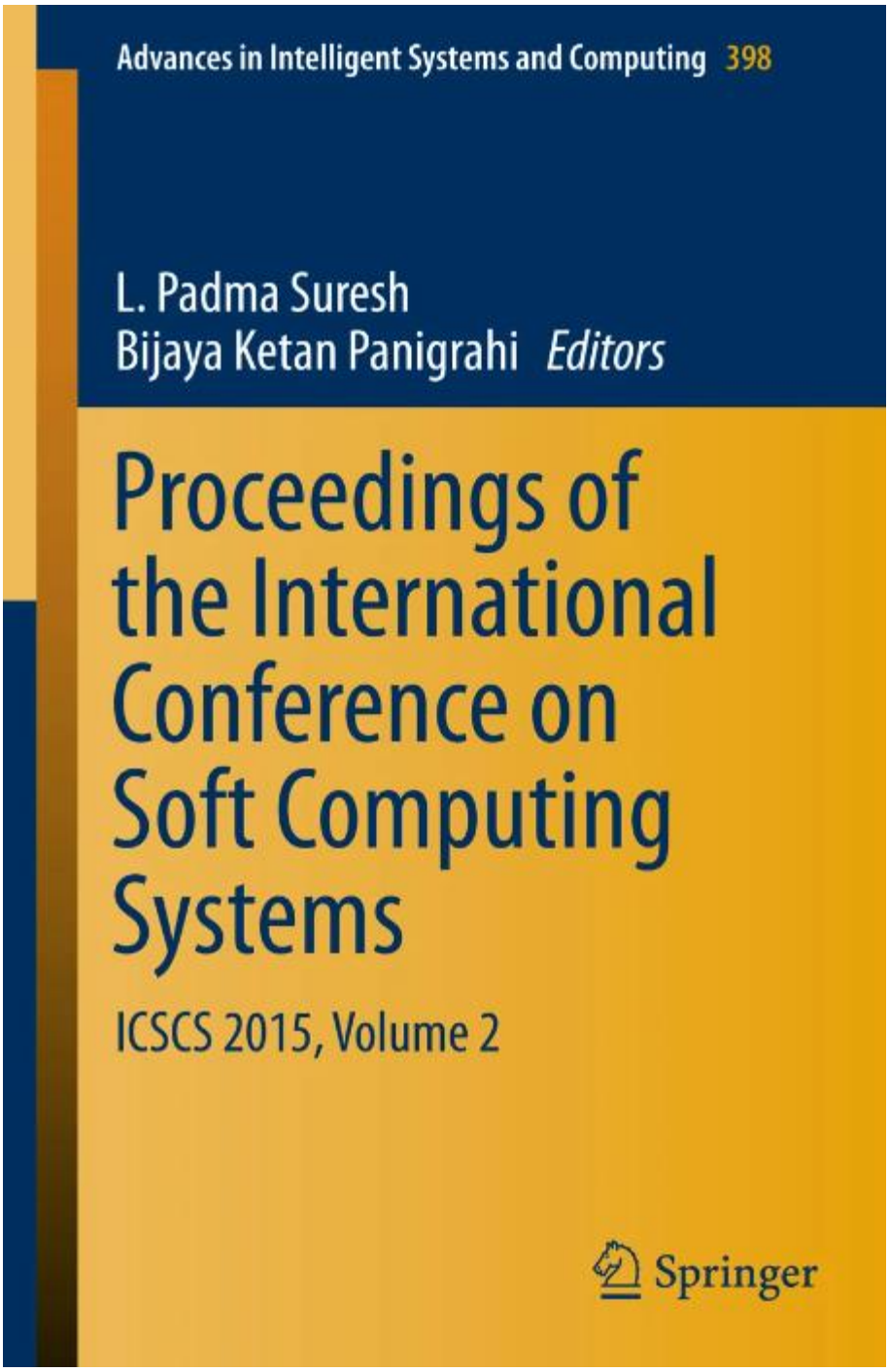
Sections Figures References

Abstract

References

Author information

Editor information


Advances in Intelligent Systems and Computing 398
L. Padma Suresh
Bijaya Ketan Panigrahi Editors
Proceedings of the International Conference on Soft Computing Systems
ICSCS 2015, Volume 2
Springer

Forecasting stock price using soft computing techniques

Publisher: IEEE Cite This PDF

N. Srinivasan ; C. Lakshmi All Authors

4 Paper Citations
172 Full Text Views

Abstract

Document Sections

I. Introduction

II. Parameter

III. Real Time Input Parameter

IV. Genetic Algorithm

Abstract:
Forecasting stock price is a important challenging task in the real world because more and more money is involved and they are affected by many social, economic, political and psychological factors. Numerous machine learning procedures have been used to fore see developments in stock cost. Machine learning classifiers include expanding the past encounters into the future actives. The proposed framework presents another hereditary calculation for forecast of monetary execution with information sets from a known source. The objective is to deliver a GA-based procedure for expectation of securities exchange execution alongside an acquainted principle classifier from an information set. We developed a genetic fuzzy based model with the ability of rule based extraction to predict next day stock price. We evaluate capacity of the proposed approach by applying it to the banking sector stocks and compare the outcome with the previous models.

Published in: 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)

Date of Conference: 23-24 March 2017

Date Added to IEEE *Xplore*: 18 January 2018

ISBN Information:

INSPEC Accession Number: 17507386

DOI: 10.1109/ICONSTEM.2017.8261274

Publisher: IEEE

Conference Location: Chennai, India

Springer Link

Search Log in



Proceedings of the Third International Conference on Computational Intelligence and Informatics pp 693–700 | Cite as

Food Consumption Monitoring and Tracking in Household Using Smart Container

Y. Bevish Jinila, V. Rajalakshmi, L. Mary Gladence & V. Maria Anu

Conference paper | First Online: 18 March 2020

545 Accesses | **3** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 1090)

Access via your institution →

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-981-15-1480-7_60
- Chapter length: 8 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

Activate Windows
Go to Settings to activate Win

Abstract

Tracking food consumption in household and monitoring the usage of items in kitchen

Abstract

Tracking food consumption in household and monitoring the usage of items in kitchen everyday is incredible. Manual intervention of this process consumes time and is annoying. This paper presents a smart container for monitoring and tracking the food items stored in household kitchen. The proposed solution serves to provide an automated response on the level of the items stored in the container and also generates an alert on its expiry. A prototype of the proposed method is developed and tested for level and expiry detection.

Keywords

Smart container **Food consumption** **Kitchen** **Household**

This is a preview of subscription content, access via your institution.

- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 181.89

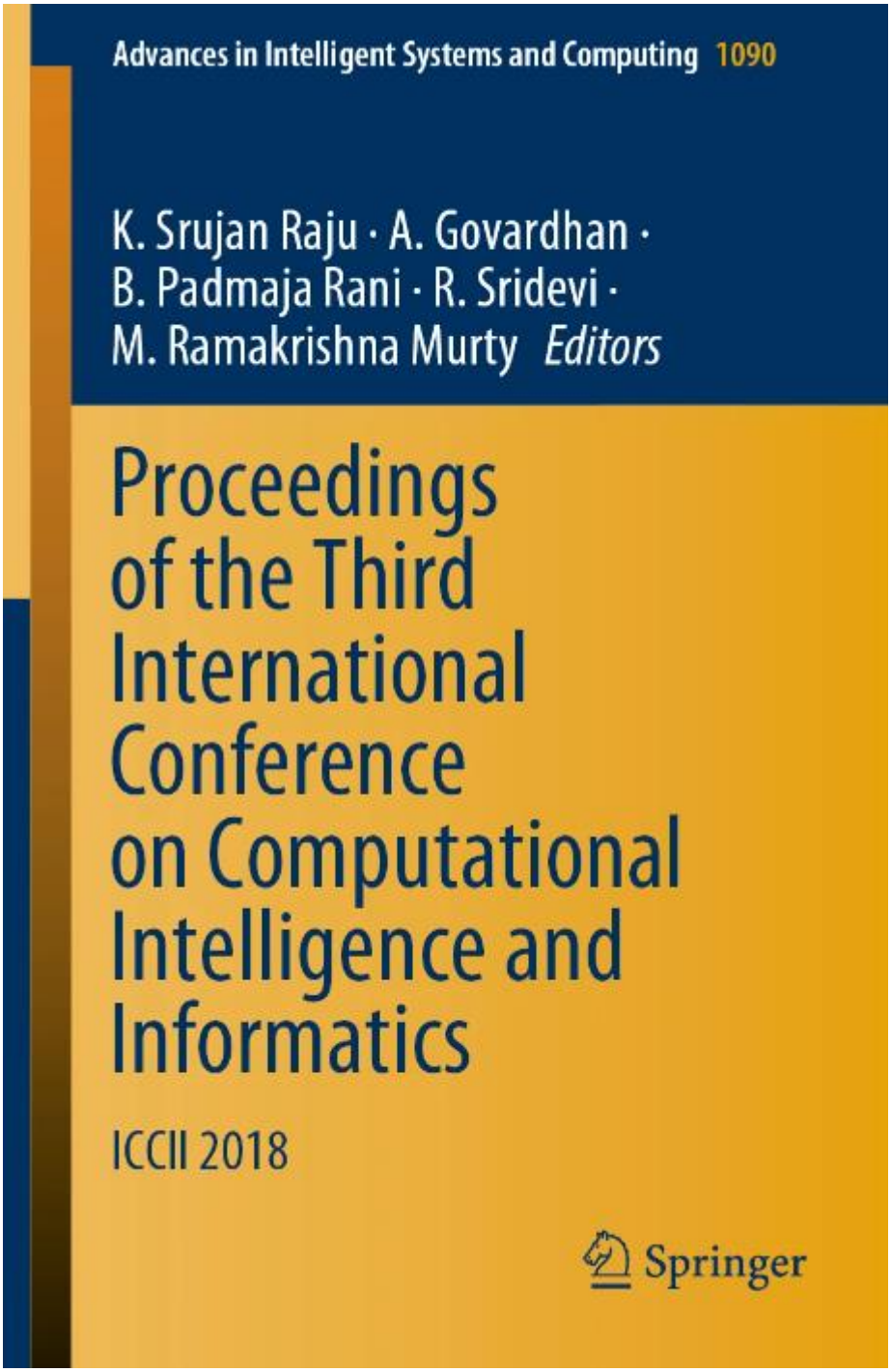
Softcover Book EUR 219.99

Learn about institutional subscriptions

Sections Figures References

Abstract

References


Advances in Intelligent Systems and Computing 1090
K. Srujan Raju · A. Govardhan ·
B. Padmaja Rani · R. Sridevi ·
M. Ramakrishna Murty Editors
Proceedings of the Third International Conference on Computational Intelligence and Informatics
ICCII 2018
Springer

Flux controlled BLDC motor for automated glass cleaning robot

Publisher: IEEE

Cite This PDF

Tushar Anand ; Sunny Kumar Kushwaha ; S. Emalda Roslin ; N. M. Nandhitha All Authors

2 Paper Citations
281 Full Text Views

Abstract

Document Sections

I. Introduction

II. Literature Review

III. Comparative Study on the Various Existing Methodologies

IV. Proposed Methodology

Abstract:

With growth of modern day infrastructure, now a days we find the peak of buildings vanishing somewhere in the clouds. All around us, there stand tall buildings having Glass as their building blocks. It also adds up to the enigma of their look. These building are mostly IT towers and Corporate Offices which need great maintenance. The cleaning of such towers is done using human resources and comprises great risk of human life, time consumption, energy consumption and is also too expensive. Moreover, cleaning is the most common work being performed in everyone's life every day. In this paper a detailed survey on various automated robots used for cleaning purpose and the principles behind the operation is done. After, surveying several advancements made in this field in the recent years we have also proposed a prototype design of an autonomous robot for glass cleaning application. The proposed model can minimize human effort and risk, energy consumption, capital and can also be efficient enough to perform. The idea is to use the principle of negative thrust-pressure to create vacuum under the robot using air blowers. The

Publisher: IEEE Cite This PDF

Tushar Anand ; Sunny Kumar Kushwaha ; S. Emalda Roslin ; N. M. Nandhitha All Authors

Department of Electronics and Control Engg., Sathyabama University, Chennai, India

2 Paper Citations | Text Views



Abstract

Document Sections

I. Introduction

II. Literature Review

III. Comparative Study on the Various Existing Methodologies

IV. Proposed Methodology

V. Conclusion

Abstract:
With growth of modern day infrastructure, now a days we find the peak of buildings vanishing somewhere in the clouds. All around us, there stand tall buildings having Glass as their building blocks. It also adds up to the enigma of their look. These building are mostly IT towers and Corporate Offices which need great maintenance. The cleaning of such towers is done using human resources and comprises great risk of human life, time consumption, energy consumption and is also too expensive. Moreover, cleaning is the most common work being performed in everyone's life every day. In this paper a detailed survey on various automated robots used for cleaning purpose and the principles behind the operation is done. After, surveying several advancements made in this field in the recent years we have also proposed a prototype design of an autonomous robot for glass cleaning application. The proposed model can minimize human effort and risk, energy consumption, capital and can also be efficient enough to perform. The idea is to use the principle of negative thrust-pressure to create vacuum under the robot using air blowers. The developed model promise less weight compared to the other models present in the literature.

Published in: 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)

Date of Conference: 23-24 March 2017

INSPEC Accession Number: 17507398

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261344

ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India

Exploring object segmentation methods in visual surveillance for human activity recognition

Publisher: IEEE Cite This PDF

Sandipkumar M Vaniya ; B. Bharathi All Authors

4 Paper Citations
231 Full Text Views

Abstract
Document Sections
I. Introduction
II. Applications
III. Object Segementation Methods

Abstract:
Human activity recognition in video camera is the prime research topic in computer vision and machine learning since last many years. Visual cameras have been used in public and private place like railway station, shopping malls, airport, offices, schools and university, etc. to recognize threat in the scene. The automated visual surveillance system will help to catch suspect in the scene, person identification in distance and re-identification, traffic management, sports, human computer interface, etc. Generally speaking, the human activity recognition in visual surveillance divided into following stages: object (human or vehicle) segmentation, feature extraction, object classification, object tracking, and activity recognition. Hence, the robust and object segmentation method in video camera is a very important phase because the rest of methods are strongly rely on it. In this paper, we have studied

Publisher: IEEE Cite This PDF

Sandipkumar M Vaniya ; B. Bharathi All Authors

Department of Computer Science & Engineering, Sathyabama University, Chennai, India

4 Paper Citations
Text Views

Abstract
Document Sections
I. Introduction
II. Applications
III. Object Segementation Methods
IV. Future Research
V. Conclusion

Abstract:
Human activity recognition in video camera is the prime research topic in computer vision and machine learning since last many years. Visual cameras have been used in public and private place like railway station, shopping malls, airport, offices, schools and university, etc. to recognize threat in the scene. The automated visual surveillance system will help to catch suspect in the scene, person identification in distance and re-identification, traffic management, sports, human computer interface, etc. Generally speaking, the human activity recognition in visual surveillance divided into following stages: object (human or vehicle) segmentation, feature extraction, object classification, object tracking, and activity recognition. Hence, the robust and object segmentation method in video camera is a very important phase because the rest of methods are strongly rely on it. In this paper, we have studied the various methods and/or algorithms of object segmentation (human). We will also discuss the strength and weakness of algorithms, complexities in activity understanding and identify the possible future research challenges.

Published in: 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)

Date of Conference: 22-24 December 2016

Date Added to IEEE *Xplore*: 26 June 2017

ISBN Information:

INSPEC Accession Number: 16980394

DOI: 10.1109/ICGTSPICC.2016.7955356

Publisher: IEEE

Conference Location: Jalgaon, India

Springer Link

Search Log in



Proceedings of the International Conference on Soft Computing Systems pp 365–376 | Cite as

Eigen-Based Indexing and Retrieval of Similar Videos

R. Sunitha & A. Mary Posonia

Conference paper | First Online: 08 December 2015

1088 Accesses | **1** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 398)

Access via your institution

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-81-322-2674-1_35
- Chapter length: 12 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 192.59

Abstract

Proficient identification of similar videos is a significant and a consequential issue in content-based video retrieval. Video search is performed on the basis of keywords and mapped to the tags associated with each video, which does not produce anticipated results. This gave rise to the concept content-based video retrieval (CBVR), which utilizes the visual features of video to store and retrieves a video effectively. Proficient retrieval of similar videos aims at creating a video by video retrieval to achieve an ontology-based retrieval of videos.

Eigen-Based Indexing and Retrieval of Similar Videos

R. Sunitha & A. Mary Posonia

Conference paper | First Online: 08 December 2015

1088 Accesses | **1** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 398)

Abstract

Proficient identification of similar videos is a significant and a consequential issue in content-based video retrieval. Video search is performed on the basis of keywords and mapped to the tags associated with each video, which does not produce anticipated results. This gave rise to the concept content-based video retrieval (CBVR), which utilizes the visual features of video to store and retrieves a video effectively. Proficient retrieval of similar videos aims at creating a video by video retrieval to achieve an ontology-based retrieval of videos.

Keywords

Video preprocessing **Key frame generation** **Pencil shaded** **Pattern extractor** **Eigen-based indexing**

Access via your institution

Chapter EUR 29.95

Price includes VAT (India)

- DOI: 10.1007/978-81-322-2674-1_35
- Chapter length: 12 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 192.59

Softcover Book EUR 229.99

Learn about institutional subscriptions

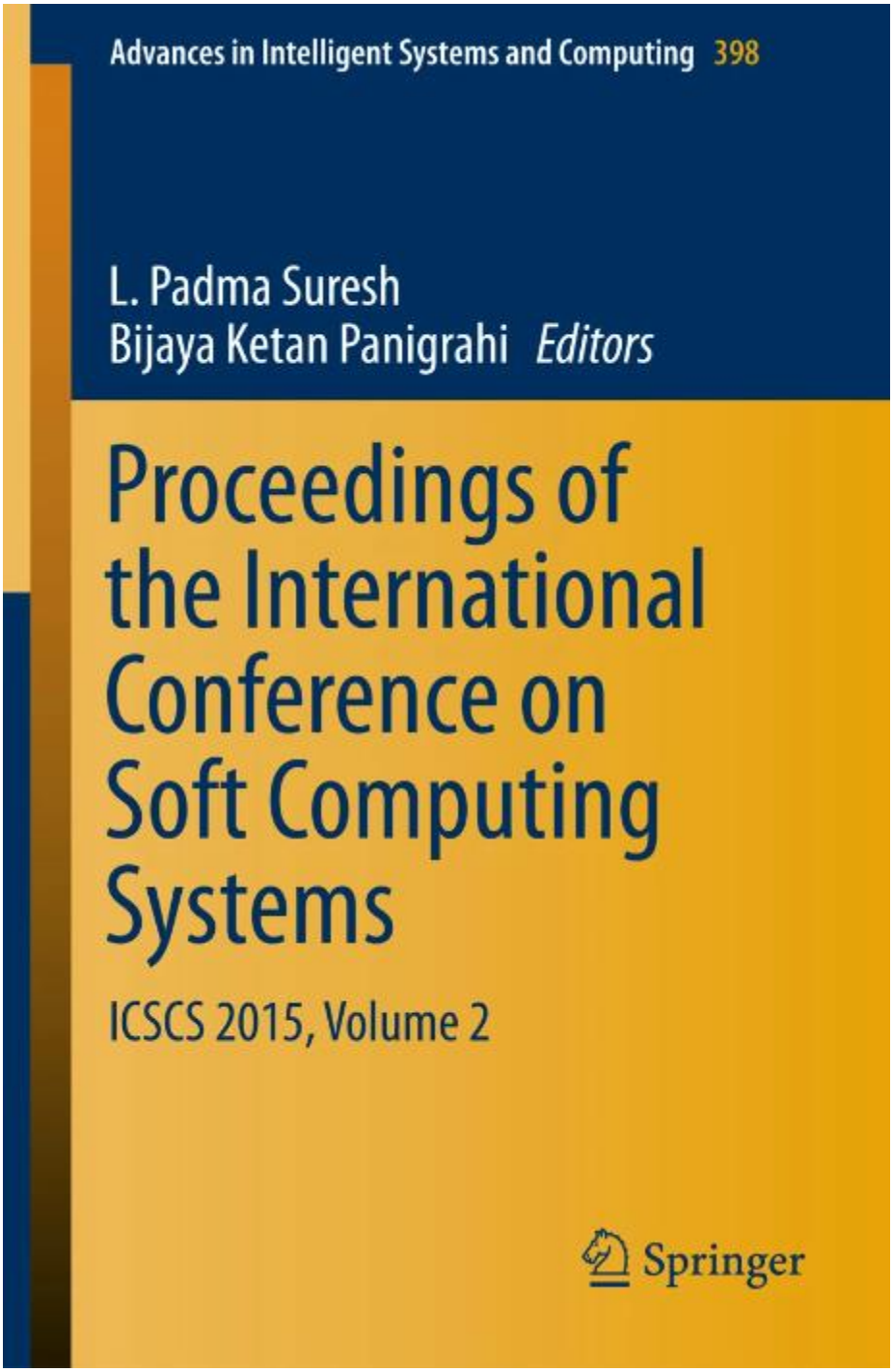
Sections Figures References

Abstract

References

Activate Windows

Go to Settings to activate Windo


Advances in Intelligent Systems and Computing 398
L. Padma Suresh
Bijaya Ketan Panigrahi Editors
Proceedings of the International Conference on Soft Computing Systems
ICSCS 2015, Volume 2
Springer

2017 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)

DYNAMIC CONTROL OF STREET LIGHTS WITH HUMAN MOBILITY, ILLUMINATION

S.Dhamodaran[1], Department of CSE, Sathyabama university, s.dhamodaran07@gmail.com

T.Kailash[2], Department of CSE Sathyabama university, kailu9505@hotmail.com

G.Tinku Srinivasan[3], Department of CSE, Sathyabama university,

Dr.M.Lakshmi[4], Faculty Of Computing, SathyabamaUniverisity, laks@icadsindia.com

I. ABSTRACT

The essential goal of the present paper is to frame a keen light post utilizing ZigBee system of gadgets. The recently proposed framework offer better effectiveness and significant power sparing can be accomplished by utilizing high proficient LED tech and furthermore productive to inform any mischances by means of remote camera to the concerned base station. Early days, the road lighting frameworks are worked physically. So the power utilization was high to the point that on the off chance that we neglected to turn off the light means it would gleam for a whole day. The prime variable is that the street side road lighting framework is a crucial angle out in the open parts. Along these lines, sun based road lighting framework ought to be composed and actualized. The photovoltaic boards are utilized to control a rechargeable battery like Pb-H+ battery, which controls a LED light. A checking station is kept in each light post and one base station as a focal center point. These checking stations contain a few modules that cooperate and exchange data to the base station through ZigBee system of gadgets. From the separate center, the relating moves will make put.

Date of Conference: 22-23 March 2017

Date Added to IEEE *Xplore*: 15 February 2018

ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD)

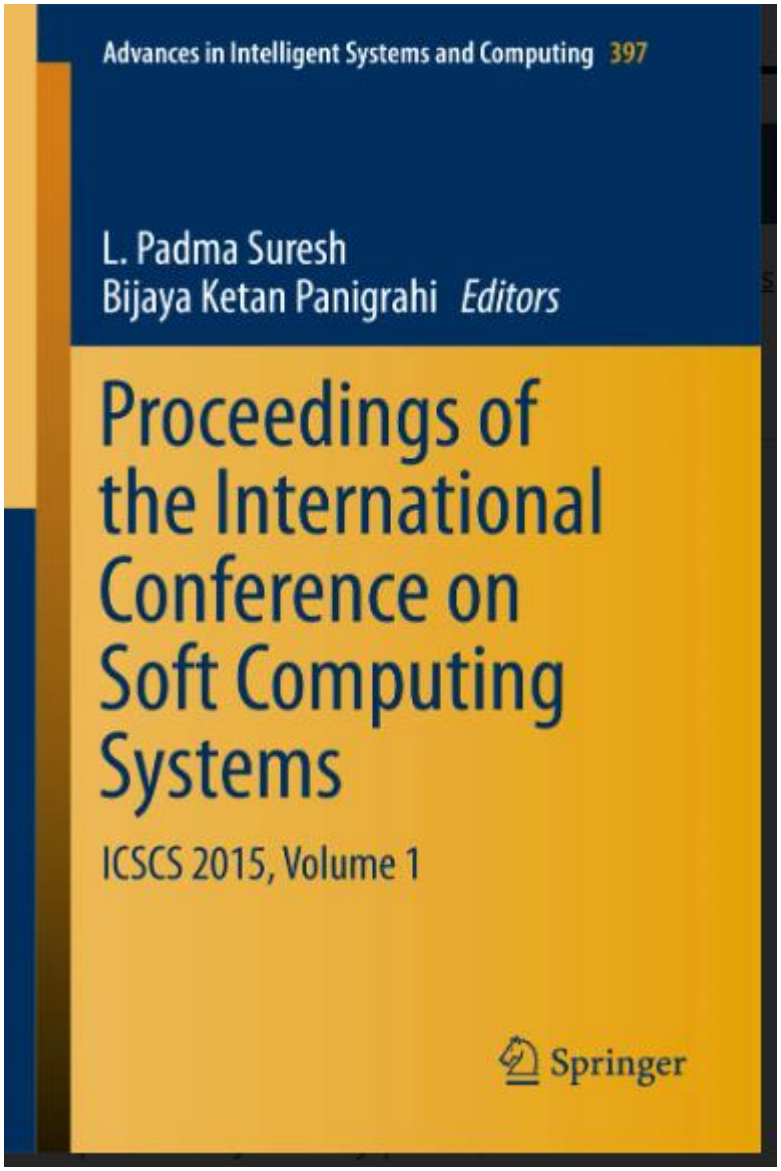
ISBN:978-1-5090-4325-5

INSPEC Accession Number: 17595380

DOI: 10.1109/ICCPEIC.2017.8290427

Publisher: IEEE

Conference Location: Melmaruvathur, India





Proceedings of the International Conference on Soft Computing Systems pp 237–249 | Cite as

Dual Converter Multimotor Drive for Hybrid Permanent Magnet Synchronous in Hybrid Electric Vehicle

G. Nagarajan ✉, C. N. Ravi, K. Vasanth, D. Godwin Immanuel & S. D. Sundarsingh Jebaseelan

Conference paper | First Online: 29 December 2015

1747 Accesses | **1** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 397)

Abstract

Electric vehicles (EV) and hybrid electric vehicles (HEV) are operated by battery power, which has to be optimized. HEV is the combination of an electric machine and an internal combustion engine (ICE) which is a promising means of reducing emissions and fuel consumption without compromising vehicle functionality and driving performance. The hybrid electric vehicle can be classified according to the way in which power is supplied to drive the train as parallel hybrid electric vehicle (PHEV), series hybrid electric vehicle (SHEV), dual-mode hybrid electric vehicle (DMHEV). In this proposed work, dual-mode hybrid electric vehicle (DMHEV) is used. One of the motors in the drive is a permanent magnet synchronous motor (PMSM). This PMSM serves two operations: As a parallel motor with an existing motor and the winding of PMSM acts as a boost inductor. A dual converter is proposed which acts as a DC-to-AC converter when PMSM is working as a motor. The same converter acts as a boost converter which uses the PMSM inductance in the boost mode. In this work, a total automobile HEV system is simulated using MATLAB. This system has an electrical subsystem which contains a generator, a battery and motors. The dual converter and PMSM are included in the electrical subsystem to improve the performance of the automobile system.

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-81-322-2671-0_23

Published	Publisher Name	Print ISBN
29 December 2015	Springer, New Delhi	978-81-322-2669-7
Online ISBN	eBook Packages	
978-81-322-2671-0	Engineering Engineering (R0)	

Dissecting different strategies from dim scale to shading

Publisher: IEEE Cite This PDF

I. Sharon ; S. SindhuParkavi ; S. Gowri ; S. Vigneshwari **All Authors**

Abstract:
Visual Cryptography (VC) is a cryptographic procedure which scrambles pictures caught utilizing different visual media to be encoded in a manner that unscrambling is done by means of sight perusing. Securing of huge information has turned into the most noteworthy and troublesome angle. This proposition is to decrease the troublesome in securing the substantial size information. The current framework utilizes dark scale picture and the bamboozling anticipation components presented in announcing CPVC plans. It either builds the pixel extension or it needs an extra validation share, which prompts to heavier weight for clients in conveying bigger/extra shares. The proposed framework addresses this issue by transforming it into coloring the picture by utilizing some steady pixels and lessens the weight for clients in conveying bigger/extra shares utilizing Visual Cryptography Schemes.

Published in: 2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)

Date of Conference: 20-21 April 2017

INSPEC Accession Number: 17279963

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICCPCT.2017.8074226

ISBN Information:

Publisher: IEEE

Conference Location: Kollam, India

Authors

I. Sharon
Dept of Information Technology, Sathyabama University, Chennai, India

S. SindhuParkavi
Dept of Information Technology, Sathyabama University, Chennai, India

S. Gowri
Dept of Information Technology, Sathyabama University, Chennai, India

S. Vigneshwari
Dept.of Computer Science Engineering, SathyabamaUniversity, Chennai, India

Design of microstrip patch antenna with multiple slots for satellite communication

Publisher: IEEE Cite This PDF

Mekala Harinath Reddy ; R. M. Joany ; G. Manikandan ; A. Sahaya Anselin Nisha **All Authors**

Abstract:
Since microstrip patch antennas are very low cost, small size and easy to fabricate they play a dominant role in wireless communication. In this paper, a microstrip patch antenna is proposed for satellite applications and it is designed at 15 GHz with multiple slots on the rectangular patch. These slots on the patch shift the resonant frequency to the lower side and also affect various antenna parameters. The antenna has a frequency bandwidth of 1.14 GHz with return loss −30.6 dB, gain 3.488 dB and directivity 3.544 dB. The antenna is designed and simulated using HIGH-FREQUENCY STRUCTURE SIMULATOR (HFSS). The antenna is designed by using RT duroid material that has relative permittivity 2.2 and loss tangent 0.0009. The results clearly explain that the proposed antenna has favorable characteristics for satellite applications.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

Date Added to IEEE *Xplore*: 08 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-3800-8
Print on Demand(PoD) ISBN:978-1-5090-3801-5

INSPEC Accession Number: 17562456

DOI: 10.1109/ICCSP.2017.8286482

Publisher: IEEE

Conference Location: Chennai, India

Authors

Mekala Harinath Reddy
Department of ECE, Sathyabama University, Chennai

R. M. Joany
Department of ECE, Sathyabama University, Chennai

G. Manikandan
Department of ECE, Sathyabama University, Chennai

A. Sahaya Anselin Nisha
Sathyabama University, Chennai, Tamil Nadu, IN

Design of cascaded wavelet filters for removing grain noise in weld radiographs

Publisher: IEEE Cite This PDF

K. Sudheera ; N.M. Nandhitha ; K. Lakshman Narayan ; K. Kunalan All Authors

Abstract:
Radiographic testing is NDT technique is used for weld quality assessment in industries. Generation of x-rays, Number of x-rays sheet and amount of energy absorbed by specimen are all random processor and hence result in noise. Specially metal alides in x-ray sheet appear as grain when digitized also the size of grain increases while zooming in digitization hence grain noise is most predominant noise in radiograph. Also an artifacts degrades the quality of x-rays. Hence it is necessary to development an efficient de-noising technique for removing noise in radiograph. In this paper a novel de-noising technique by using spatial and spectral domain and optimization of wavelets is done. Daubechies wavelet 2 provides better de-noising than the other techniques.

Published in: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS)

Date of Conference: 15-16 June 2017

INSPEC Accession Number: 17487168

Date Added to IEEE *Xplore*: 11 January 2018

DOI: 10.1109/ICCONS.2017.8250674

ISBN Information:
Electronic ISBN:978-1-5386-2745-7
CD:978-1-5386-2744-0
Print on Demand(PoD) ISBN:978-1-5386-3901-6

Publisher: IEEE

Conference Location: Madurai, India

Authors

K. Sudheera
School of EEE Sathyabama University, Chennai, India

N.M. Nandhitha
School of EEE Sathyabama University, Chennai, India

K. Lakshman Narayan
B.E ECE, Chennai, India

K. Kunalan
School of EEE Sathyabama University, Chennai, India

Design and implementation of security system for smart home

Publisher: IEEE

Cite This

PDF

M Akhil Raja ; G Rakesh Reddy ; Ajitha **All Authors**

Abstract:

Security is becoming a major issue everywhere. Home Security is becoming necessary nowadays as the possibilities of intrusion is increasing day by day. Safety from theft, leaking of raw gas and fire are the necessary requirements for home security system. However, the GSM (global system for mobile communication) based security systems provides enhanced security as whenever a signal from sensor occurs, a text message is sent to the desired number to take necessary actions. Here we are using an antitheft reporting system which will report the owner by ringing alarm and send a text message. Also for the safety system in case of fire or gas leakage, it will notify the owner by sending an SMS and also by ringing an alarm.

Published in: 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)

Date of Conference: 16-18 February 2017

Date Added to IEEE *Xplore*: 14 December 2017

ISBN Information:

Electronic ISBN:978-1-5090-3378-2

Print on Demand(PoD) ISBN:978-1-5090-3379-9

INSPEC Accession Number: 17429841

DOI: 10.1109/ICAMMAET.2017.8186705

Publisher: IEEE

Conference Location: Chennai, India

Authors

M Akhil Raja

Department Of IT, Sathyabama University, Chennai, India

G Rakesh Reddy

Department Of IT, Sathyabama University, Chennai, India

Ajitha

Department Of IT, Sathyabama University, Chennai, India

IOP Conf. Series: Materials Science and Engineering **590** (2019) 012047 doi:10.1088/1757-899X/590/1/012047

DESIGN AND IMPLEMENTATION OF C III FILTER FOR MEASURING DISRUPTION IN ADITYA TOKAMAK

L. Jegan Antony Marcilin[1]
Assistant Professor, Department of ECE,
Sathyabama Institute of Science and Technology,
Jeppiaar Nagar, Rajiv Gandhi Salai,,Chennai-India
jegan25@hotmail.com
N.M.Nandhitha[2]
Professor & Dean, Department of ECE,
Sathyabama Institute of Science and Technology,
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai - India
nandhi_n_m@yahoo.co.in
Gayathri Sam[3,]
Final Year UG student, Department of ECE,
Sathyabama Institute of Science and Technology
Jeppiaar Nagar, Rajiv Gandhi Salai,,Chennai-India
vinisam05@gmail.com
M.Gladies Jaya Chitra[3]
Final Year UG student, Department of ECE,
Sathyabama Institute of Science and Technology
Jeppiaar Nagar, Rajiv Gandhi Salai,,Chennai-India
mgladiesthangaraj@gmail.com

Abstract- The main objective of this project is to filter carbon gas among many other gases emitted from the plasma in Aditya Tokamak. The aim is to measure the disruptions released by plasma due to some unpredictable conditions. C3 filter detects carbon gases from the various gases emitted from the plasma in Aditya Tokomak. The filter is to be designed with 5-6 layers of the material chosen. Then the signal segmentation for transmittance vs. wavelength is to be achieved using MATLAB software.

Keywords: Aditya Tokamak, C III Filter, Matlab.

Design and analysis of spatial-temporal model using hydrological techniques

Publisher: IEEE Cite This PDF

S. Dhamodaran ; M. Lakshmi All Authors

Abstract:
A spatiotemporal question can be characterized as a protest that has no less than one spatial and one fleeting property. The spatial properties are the area and geometry of the question. The transient property is a timestamp or the time interim for which the protest is legitimate. spatiotemporal protest, for the most part contains spatial, worldly and topical or non-spatial attributes. Spatiotemporal informational indexes basically catch changing estimations of spatial and topical characteristics over a timeframe. In the crieteria to evalute the different procedure and adequately plan the zones to actualize the quality demonstrating utilizing spacial worldly parameters. To plan the likelihood to apply in the watershed territory, Physically based model for overflow/surge analysis. Develop and outline the calculated model for disaggregation of every day precipitation, spatial consistance, atmosphere district for various rainstorm are dissected.

Published in: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date of Conference: 22-23 March 2017

Date Added to IEEE *Xplore*: 15 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-4324-8
Print on Demand(PoD) ISBN:978-1-5090-4325-5

INSPEC Accession Number: 17580333

DOI: 10.1109/ICCPEIC.2017.8290333

Publisher: IEEE

Conference Location: Melmaruvathur, India

Authors

S. Dhamodaran
Faculty Of Computing, Sathyabama University, Chennai

M. Lakshmi
Faculty Of Computing, Sathyabama University, Chennai

Data protection and outsourcing in cloud with Linear programming and image based OTP

Publisher: IEEE Cite This PDF

Abstract:
Distributed computing has turned into a reality which is capable of empowering overall population and associations to make utilization of colossal computational assets without capital interest in pay per utilize form. This new worldview permits clients who have so much computational assets to outsource their costly workloads related to computation which is in return connected to cloud as well as obtain advantages of its stockpiling, servers, organizing and a few offices. In this manner distributed computing has endless conceivable outcomes as a result a person could realize a global standard foundation and offices in her own PC sans venture. In spite of the fact that cloud specialist organizations deal with security client have their own worries regarding insurance of their touchy information. Along these lines, it is groundbreaking to have secure system to ensure information of clients furthermore guarantee the information respectability. This will empower clients to outsource their works courageously. This proposed framework concentrates on the Linear programming calculations that happen over cloud with excellent security. Keeping in mind the eventually goal is to accomplish reasonable effectiveness, our system configuration unequivocally separate the Linear Programming (LP) computation outsourcing into open LP solvers functioning on the cloud as well as private LP parameters possessed by the client. The subsequent adaptability permits us to observe fitting security tradeoff by means of larger amount deliberation LP calculations than the general circuit representation. Furthermore for keep up information protection in cloud we utilize blow fish and picture based one-time secret password. It is likewise found that the outcome confirmation of the framework is computationally proficient does not acquire extra charge.

Published in: 2017 International Conference on Information Communication and Embedded Systems (ICICES)

Date of Conference: 23-24 February 2017

INSPEC Accession Number: 17261627

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070745

ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Authors

Rahul Kumar
Department of computer science and engineering, Sathyabama University, Chennai, India

A. Pravin
Department of computer science and engineering, Sathyabama University, Chennai, India

Comprehensive Review of Single Switch DC-DC Converters for Voltage Lift in RES Application

Publisher: IEEE Cite This PDF

Abstract:
In current epoch of increased RES applications in power engineering, there is a wide-spread research on DC-DC converter topologies to strap renewable energy source (RES) to load. In the context of boosting the voltage output of RES using DC-DC converter, many DC-DC converter topologies are discussed in the literature. This paper intends to present a comprehensive review of three coupled inductor based non-isolated type DC-DC converter topologies operated with single power switch for voltage lift in RES applications. Voltage gain formulations are presented for reviewed topologies of non-isolated DC-DC converters operating with single power switch for voltage lift in RES applications. Simulation results are presented showing voltage across switch, current through inductor, voltage across diode. DC-DC converter topologies are developed and the results are obtained using MATLAB/SIMULINK software.

Published in: 2018 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)

Date of Conference: 22-23 February 2018

Date Added to IEEE *Xplore*: 04 November 2018

ISBN Information:
Electronic ISBN:978-1-5386-3817-0
Print on Demand(PoD) ISBN:978-1-5386-3818-7

INSPEC Accession Number: 18245428

DOI: 10.1109/ICPECTS.2018.8521598

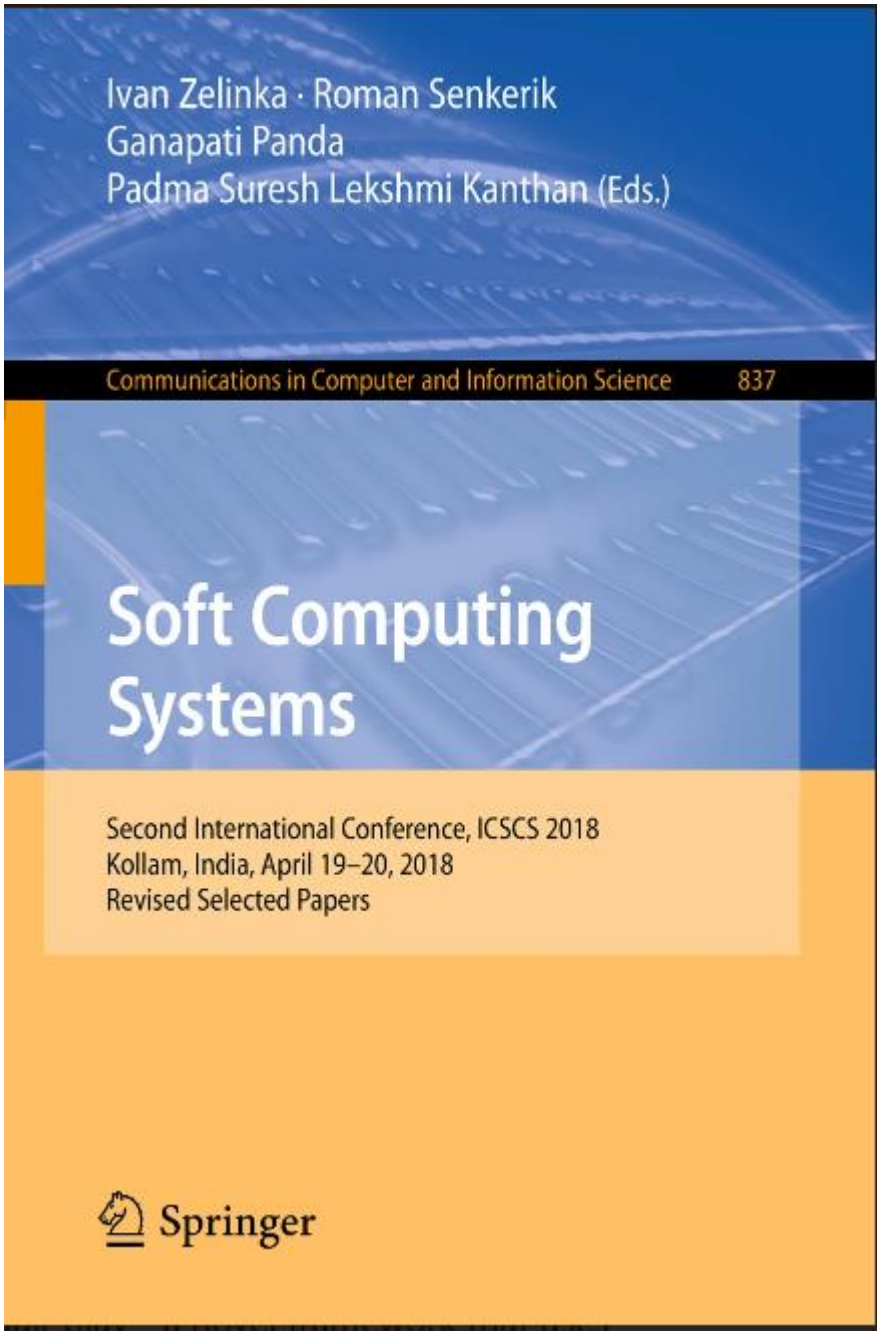
Publisher: IEEE

Conference Location: Chennai, India

Authors

Ravi Kiran Dasari
Department of Electrical and Electronics Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, India

D. Godwin Immanuel
Department of Electrical and Electronics Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, India



Comprehensive Behaviour of Malware Detection Using the Machine Learning Classifier

P. Asha ✉, T. Lahari & B. Kavya

Conference paper | First Online: 25 September 2018

1325 Accesses | **2** Citations

Part of the Communications in Computer and Information Science book series (CCIS,volume 837)

Abstract

Everyone is using mobile phone and android markets like Google play and the model they offer to certain apps make the Google play market for their false and malware. Some developers use different techniques to increase their rank, increasing popularity through fake reviews, installation accounts and introduce malware to mobile phones. Application developers use various advertising campaigns showing their popularity as the highest ranking application. They manipulate ranking on the chart. In the past they worked on application permission and authorization. In this we propose a fair play - a novel framework that uses traces left to find rank misrepresentation and applications subjected to malware. Fair play uses semantic and behavioural signs gathered from Google play information.

Cite this paper

Asha, P., Lahari, T., Kavya, B. (2018). Comprehensive Behaviour of Malware Detection Using the Machine Learning Classifier. In: Zelinka, I., Senkerik, R., Panda, G., Lekshmi Kanthan, P. (eds) Soft Computing Systems. ICSCS 2018. Communications in Computer and Information Science, vol 837. Springer, Singapore. https://doi.org/10.1007/978-981-13-1936-5_49

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-1936-5_49

Published	Publisher Name	Print ISBN
25 September 2018	Springer, Singapore	978-981-13-1935-8
Online ISBN	eBook Packages	
978-981-13-1936-5	Computer Science Computer Science (R0)	

Complexities in developing multilingual on-line courses in the Indian context

Publisher: IEEE Cite This PDF

Kolla Bhanu Prakash ; Arun RajaRaman ; M Lakshmi All Authors

Abstract:
Two major technological developments — cellular and internet communications — that have significantly changed the outlook of society throughout the world and more so in the Indian context. Usage and interaction through cell-phones and internet are so popular in different parts of India without language or dialect barrier. Education is one discipline where Indian input has been phenomenal both in terms of students and institutions and soon the switch over to online education and learning will take place to cater to different levels and 24/7 availability. The present study looks at issues in having a common web-based content for developing the courses. Towards this object and subject issues are taken and based on a generic pixel based approach, a common reference content prediction is looked at. Case studies on 'circle' — object -,'internet' — a subject and 'magnetism' — a part of science and physics demonstrate applicability and versatility in handling different document types.

Published in: 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC)

Date of Conference: 23-25 March 2017

Date Added to IEEE *Xplore*: 19 October 2017

ISBN Information:
Electronic ISBN:978-1-5090-6400-7
USB ISBN:978-1-5090-6399-4
Print on Demand(PoD) ISBN:978-1-5090-6401-4

INSPEC Accession Number: 17261541

DOI: 10.1109/ICBDACI.2017.8070860

Publisher: IEEE

Conference Location: Chirala, Andhra Pradesh, India

Authors

Kolla Bhanu Prakash
Chirala Engineering College, NLP Research Lab, Chirala, Andhra Pradesh, India

Arun RajaRaman
Indian Institute of Technology Madras, Chennai, Tamil Nadu, IN

M Lakshmi
Dean School of Computing, Sathyabama University, Chennai, India

Comparative Study of Syntactic Search Engine and Semantic Search Engine: A Survey

Publisher: IEEE Cite This PDF

A.C. Santha Sheela ; C. Jayakumar All Authors

Abstract:
Search engine symbolizes an extremely powerful and valuable tool for fetching any sort of information from Internet. There has been numerous researches carried on search engines techniques, the major ones are syntactic and semantic. Referring to the Syntactic web, the results obtained are purely as per the keyword match. That is the query outputs numerous web pages against the keyword match that may not even be relevant or meaningful. Whereas, unlike the syntactic web, the semantic web is a revised or upgraded version of the web which produces quiet meaningful and specific output as it has the potential to comprehend the query effectively. Few examples of Semantic based search engines include Kosmix, Hakia, Cognition, Swoogle and Lexxe. Whereas syntactic based search engines are Google, Yahoo, Ask. The work performs a comparison amidst the performance of semantic and syntactic based search engine and evaluates them by employing certain queries.

Published in: 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)

Date of Conference: 14-15 March 2019

Date Added to IEEE *Xplore*: 05 December 2019

ISBN Information:
Electronic ISBN:978-1-7281-1599-3
Print on Demand(PoD) ISBN:978-1-7281-1600-6

INSPEC Accession Number: 19212997

DOI: 10.1109/ICONSTEM.2019.8918837

Publisher: IEEE

Conference Location: Chennai, India

Authors

A.C. Santha Sheela
CSE Dept, Sathyabama Institute of Science & Technology, Chennai, India

C. Jayakumar
CSE Dept, Sri Venkateswara College of Engineering, Chennai, India

Cognitive computing techniques based rainfall prediction — A study

Publisher: IEEE Cite This PDF

J Refonaa ; M Lakshmi All Authors

Abstract:
The Primary source of agriculture in India is rainfall. Even though India has its own surplus fertile land, agriculture struggles a lot due to lack of efficient predictable rainfall technique which in turn leads to uncountable farmer's suicide activities. Nowadays, it becomes a back-breaking problem giving way to scarcity of food storage. Even Chennai faces a massive hit in 2015 december and 2016 december and flattens many residential areas and people started moving from their natives and starve for food with their small kids and old age people. So In this paper, cognitive computing prediction model methodologies, their outcomes and limitations are studied for the development of new hybrid cognitive model that supports the accuracy of prediction rate.

Published in: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date of Conference: 22-23 March 2017

Date Added to IEEE *Xplore*: 15 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-4324-8
Print on Demand(PoD) ISBN:978-1-5090-4325-5

INSPEC Accession Number: 17580294

DOI: 10.1109/ICCPEIC.2017.8290354

Publisher: IEEE

Conference Location: Melmaruvathur, India

Authors

J Refonaa
Research Scholar: School of Computing, Sathyabama university, Chennai, Tamil Nadu

M Lakshmi
Prof & Dean: School of Computing, Sathyabama University, Chennai, Tamil Nadu



Advances in Big Data and Cloud Computing pp 539–546 | Cite as

Cloud-Based Scheme for Household Garbage Collection in Urban Areas

Y. Bevish Jinila, Md. Shahzad Alam & Prabhu Dayal Singh

Conference paper | First Online: 12 December 2018

746 Accesses | **4** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 750)

Abstract

Technical advancements have rapidly increased the population in urban areas. People residing in urban areas are scheduled with regular activities. The time-spend for managing the household cleanliness and garbage disposal on daily basis incurs extra overhead. A key challenge is to provide a proper garbage disposal for houses. In the present scenario, residents have to carefully monitor the garbage bins each day and dispose it when it is full. So, the problem is overflowing of garbage results in bad odour at home and surroundings. This limitation can be addressed by replacing the garbage bins with smart bins and connecting them to a cloud server. The processing entity in the cloud server further runs a shortest path route optimization algorithm and presents the shortest path for every 24 h basis. The employee can utilize this information for the collection of garbage from houses. The prototype of the proposed method of garbage collection is developed.

Keywords

Garbage collection **Smart bin** **Shortest path routing** **Cloud computing** **Waste management**

Cite this paper

Bevish Jinila, Y., Shahzad Alam, M., Dayal Singh, P. (2019). Cloud-Based Scheme for Household Garbage Collection in Urban Areas. In: Peter, J., Alavi, A., Javadi, B. (eds) Advances in Big Data and Cloud Computing. Advances in Intelligent Systems and Computing, vol 750. Springer, Singapore. https://doi.org/10.1007/978-981-13-1882-5_47

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-1882-5_47

Published	Publisher Name	Print ISBN
12 December 2018	Springer, Singapore	978-981-13-1881-8
Online ISBN	eBook Packages	
978-981-13-1882-5	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)	

Closed loop speed controlled two stage PFC converter fed PMSM drive

Publisher: IEEE Cite This PDF

Janakiraman ; Ramesh Babu **All Authors**

Abstract:
This paper presents the closed loop controlled two stage PFC (Power Factor Correction) converter fed PMSM (Permanent Magnet Synchronous Motor) drive. The proposed PFC converters consist of uncontrolled rectifier, PFC converter and DC converter. The π filter is connected at the output of converter to reduce output ripple of DC converter. The performance of the proposed converter is verified by interfacing inverter and PMSM drive. This work deals with comparison of PI & FL controlled two stage PFC converter fed PMSM drive system. The steady state performance parameters like output voltage ripple, speed, power factor, THD are analyzed with PI (Proportional Integrative) and Fuzzy Logic Controller(FLC). The Transient parameters such as rise time, peak time, settling time, and steady state error also compared for both controller. The Closed loop PI & FL controlled two stage PFC converter Fed PMSM drive systems are modeled and simulated.

Published in: 2017 Innovations in Power and Advanced Computing Technologies (i-PACT)

Date of Conference: 21-22 April 2017

INSPEC Accession Number: 17470946

Date Added to IEEE *Xplore*: 04 January 2018

DOI: 10.1109/IPACT.2017.8245117

ISBN Information:

Electronic ISBN:978-1-5090-5682-8

Print on Demand(PoD) ISBN:978-1-5090-5683-5

Publisher: IEEE

Conference Location: Vellore, India

Authors

Janakiraman
EEE Dept, Sathyabama University, Chennai, India

Ramesh Babu
Sathyabama University, Chennai, Tamil Nadu, IN

Classification and Prediction of Breast Cancer using Linear Regression, Decision Tree and Random Forest

Publisher: IEEE Cite This PDF

Abstract:
Breast Cancer is one of a major issue that some of the women are facing today. Earlier detection of cancer by performing detailed analysis based on the existing records which may assist the physicians in providing a better treatment to their patients. Data to analyze and predict the breast cancer are obtained from UCI Machine Learning Repository (Wisconsin Breast Cancer). The main objective is to classify whether the type of cancer is benign or malignant. Based on the available data set and the patient record, whether the disease is curable or non-curable is predicted. Thus the success rate of classification is 84.14% and the prediction percentage is 88.14%.

Published in: 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)

Date of Conference: 08-09 September 2017

Date Added to IEEE *Xplore*: 06 September 2018

ISBN Information:
Electronic ISBN:978-1-5386-3243-7
Print ISBN:978-1-5386-3242-0
CD:978-1-5386-3240-6
DVD ISBN:978-1-5386-3241-3
Print on Demand(PoD) ISBN:978-1-5386-3244-4

INSPEC Accession Number: 18062687

DOI: 10.1109/CTCEEC.2017.8455058

Publisher: IEEE

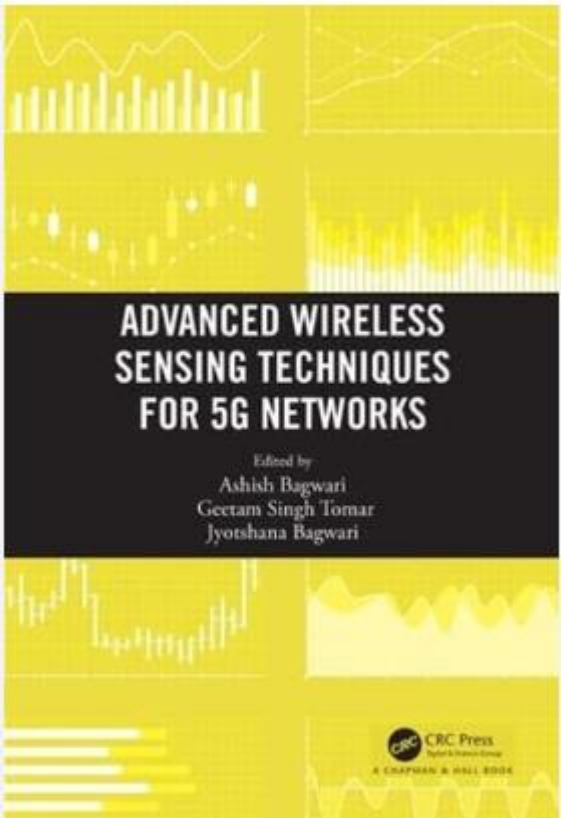
Conference Location: Mysore, India

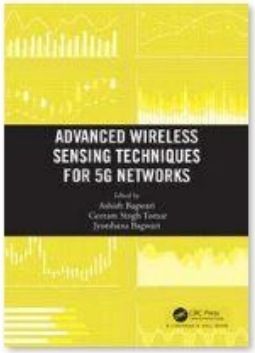
Authors

S. Murugan
Department of Computer Science and Engineering, Sathyabama University, Chennai, India

B. Muthu Kumar
Department of Computer Science and Engineering, Sathyabama University, Chennai, India

S. Amudha
Department of Computer Science and Engineering, Sathyabama University, Chennai, India





Chapter

Challenges Confronting the Next Generation Wireless Networks

By R. Narmadha, A. Jayanthiladevi

Book Advanced Wireless Sensing Techniques for 5G Networks

Edition	1st Edition
First Published	2018
Imprint	Chapman and Hall/CRC
Pages	16
eBook ISBN	9781351021746

ABSTRACT

Actualizing a practical internetwork [1] is no basic task. Many difficulties must be confronted, particularly in the regions of availability, unwavering quality, organization of administration, and adaptability. Every area is geared toward the build-up of a proficient and compelling internetwork. The main obstacle confronted when associating different frameworks is to help create correspondence between dissimilar techniques. Distinctive locales, for instance, may utilize diverse types of media or they may work at different rates. Another basic area, dependable administration, must be kept up in any internetwork. Singular clients and whole associations rely upon predictable, dependable access in order to organize assets. In addition, various administrations [2,3] must be brought together to assist each other and investigate the capabilities in an internetwork. Arrangement, security, execution, and other issues must be satisfactorily handled for the internetwork to work easily. Adaptability, the last concern, is important for arranging the development of the network, new applications, and the various administrative associations among different components.

Chapter 9 | 19 pages

Investigation of Energy Efficient Routing Protocols in Wireless Sensor Networks with Variant Battery Models

By Vinod Kumar Verma, Surinder Singh, N.P. Pathak

Abstract

GET ACCESS

Part Part V | 2 pages

Challenges in 5G Networks

Chapter 10 | 16 pages

Challenges Confronting the Next Generation Wireless Networks

By R. Narmadha, A. Jayanthiladevi

Abstract

GET ACCESS

Chapter 11 | 11 pages

Antennas in Cognitive Radio Systems for 5G Networks: Challenges and Review

By M. Hassanein Rabah, Divitha Seetharamdoo

Abstract

GET ACCESS

Brain tumor detection based on watershed transformation

Publisher: IEEE Cite This PDF

K. Ramya ; L. K. Joshila Grace **All Authors**

Abstract:
In this paper, a watershed transformation technique is used with gradient magnitude with morphological open image and two important features is used as foreground and background to identify the tumor. First the Magnetic Resonance imaging (MRI) Scan of tumor is given as an input and it undergoes into watershed technique which is a topological boundary dividing into two adjacent brain cells. With the gradient magnitude for segmentation technique the rate of inclination or declination of a tumor will be identified. To identify the foreground of the tumor, open the image morphological, thus it acquires clear idea about how the particular tumor will be closer to normal cells. With the background marker, the invisible tumor will be identified using threshold value. In the segmentation output finally, the intensity, size, shape of the tumor in the brain is displayed and can be analysis.

Published in: 2016 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2016

Date Added to IEEE *Xplore*: 24 November 2016

ISBN Information:
Electronic ISBN:978-1-5090-0396-9
CD:978-1-5090-0395-2
Print on Demand(PoD) ISBN:978-1-4673-8549-7

INSPEC Accession Number: 16498129

DOI: 10.1109/ICCSP.2016.7754379

Publisher: IEEE

Conference Location: Melmaruvathur, India

Authors

K. Ramya
Sathyabama University, Chennai, Tamil Nadu, IN

L. K. Joshila Grace
Sathyabama University, Chennai, Tamil Nadu, IN

Blood donation and life saver app

Publisher: IEEE

Cite This PDF

M R Annish Brislin ; J Albert Mayan ; R Aroul Canessane ; M R Anish Hamlin **All Authors**

Abstract:
Blood donation saves lives everyday at various situations. A blood transfusion may give them the energy to spend time with family and friends. Blood cannot be manufactured it can only comes as a gift from people. One person only allowed 6 pints of blood donation. One pint of blood can save upto 33 lives The number of blood donor is very less when compared with other countries. Here we propose a new and efficient way to overcome such outline. When we just touch the button donor the App will be ask you to enter an individual's details like name, phone number, age, weight, date of birth, blood group, address etc. At the emergency time of blood needed we can check for blood donor nearby by using GPS. Once the app user enter the blood group which he/she needed it will automatically show the donor nearby and send an alert message to the donor. In case if the first donor is not available it will automatically search the next donor which is present in queue. If the donor accept the request then an one time password(OTP) will be send to the donor to verify. Blood donation app provider list of donor in your city/area. Once the donor donate the blood it will automatically remove the donor detail for next three months.

Published in: 2017 2nd International Conference on Communication and Electronics Systems (ICCES)

Date of Conference: 19-20 October 2017

Date Added to IEEE *Xplore*: 22 March 2018

ISBN Information:
Electronic ISBN:978-1-5090-5013-0
DVD ISBN:978-1-5090-5012-3
Print on Demand(PoD) ISBN:978-1-5090-5014-7

INSPEC Accession Number: 17650561

DOI: 10.1109/CESYS.2017.8321318

Publisher: IEEE

Conference Location: Coimbatore, India

Authors

M R Annish Brislin
Department of Electronics and Communication, Jeppiaar SRR Engineering College, Chennai

J Albert Mayan
Department of Computer Science and Engineering, Sathyabama University, Chennai

R Aroul Canessane
Department of Computer Science and Engineering, Sathyabama University, Chennai

M R Anish Hamlin
Department of Computer Science and Engineering, Sathyabama University, Chennai

Bandwidth enhancement of microstrip patch antenna using parasitic patch

Publisher: IEEE Cite This PDF

Mekala Harinath Reddy ; R.M. Joany ; M. Jayasaichandra Reddy ; M. Sugadev ; E. Logashanmugam All Authors

Abstract:
Microstrip patch antenna finds a better choice in wireless communications because of their less cost, light weight and ease of fabrication. This paper proposes a novel design for the bandwidth enhancement of microstrip patch antenna with parasitic patch placed adjacent to the radiating patch along with the slots. The proposed structure is designed at 8.5 GHz using FR4-epoxy as substrate material that has relative permittivity 4.4 and loss tangent 0.02. HFSS (High Frequency Structure Simulator) software is used for simulating the proposed design. This antenna obtained the return loss -41.4.30 dB, gain 5.3542 dB, directivity 5.4612 dB and VSWR 1.0172. Comparative analysis with various slots on both main patch and parasitic patch has been done for various antenna parameters like; return loss, gain, directivity and VSWR. The results clearly explain that the proposed design has favourable characteristics suitable for X-band applications.

Published in: 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)

Date of Conference: 02-04 August 2017

Date Added to IEEE *Xplore*: 30 October 2017

ISBN Information:
Electronic ISBN:978-1-5090-5905-8
Print ISBN:978-1-5090-5918-8
CD:978-1-5090-5904-1
Print on Demand(PoD) ISBN:978-1-5090-5906-5

INSPEC Accession Number: 17316085

DOI: 10.1109/ICSTM.2017.8089172

Publisher: IEEE

Conference Location: Chennai, India

Authors

Mekala Harinath Reddy
Department of ECE, Sathyabama University, Chennai, India

R.M. Joany
Department of ECE, Sathyabama University, Chennai, India

M. Jayasaichandra Reddy
Department of ECE, Sathyabama University, Chennai, India

M. Sugadev
Department of ECE, Sathyabama University, Chennai, India

E. Logashanmugam
Department of ECE, Sathyabama University, Chennai, India

Automation of street light for smart city

Publisher: IEEE

Cite This

PDF

M. Revathy ; S. Ramya ; R. Sathiyavathi ; B. Bharathi ; V. Maria Anu **All Authors**

Abstract:
Energy consumption in metropolitan cities is increasing day by day. In every city, considerable amount of electricity is being used for the purpose of street lighting system. Some areas of the city may have low frequency of passerby, but it's observed that the amount of energy the street lights consume in these areas is same as that of areas with high frequency passerby. As a result enormous amount of energy is wasted without being used. In the proposed system, high intensity discharge lamps are replaced by LED's which can alter its intensity based on the need. Movement of vehicles is sensed using LDR (Light Dependent Resistor) and the intensity of the street light is reduced when not in use. The system also detects fault in the system and indicates it to the base station using GSM (Global System for Mobile communication) technology by sending SMS (short message service).

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

Date Added to IEEE *Xplore*: 08 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-3800-8
Print on Demand(PoD) ISBN:978-1-5090-3801-5

INSPEC Accession Number: 17562543

DOI: 10.1109/ICCSP.2017.8286503

Publisher: IEEE

Conference Location: Chennai, India

Authors

M. Revathy
School of Computing, Sathyabama University, Chennai, India

S. Ramya
School of Computing, Sathyabama University, Chennai, India

R. Sathiyavathi
School of Computing, Sathyabama University, Chennai, India

B. Bharathi
School of Computing, Sathyabama University, Chennai, India

V. Maria Anu
School of Computing, Sathyabama University, Chennai, India

Automatic traffic control system using PCA based approach

Publisher: IEEE Cite This PDF

D Usha Nandini ; M Saravanan ; J Albert Mayan ; Murari Devakannan Kamalesh ; K Mohana Prasad All Authors

Abstract:
In recent years, communication in wireless is well improved technology for moving the data like text, image, video, and audio. But in WN (wireless network), overload is a important issues of packet loss and traffic, and high excellence video and bandwidth limit are important network surplus thread. To defeat these issues, Shih-Chia Huang Bo-Hao suggested PCA-related RBF network method for moving object identification (motion detection). This method first changes the video bit rate as per the level of bandwidth. If bandwidth level is more, then the video bit rate would be more, if bandwidth level is minimum then the moving bit rate of video would be more. Also this method identifies the moving object in low video bit rate. We apply this method on automatic traffic control system. In existing there is no automatic system for monitor the traffic system. These above approaches are useful in identifying the traffic level and to control the traffic automatically

Published in: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)

Date of Conference: 01-02 August 2017

Date Added to IEEE *Xplore*: 21 June 2018

ISBN Information:
Electronic ISBN:978-1-5386-1887-5
Print on Demand(PoD) ISBN:978-1-5386-1888-2

INSPEC Accession Number: 17859616

DOI: 10.1109/ICECDS.2017.8389877

Publisher: IEEE

Conference Location: Chennai, India

Authors

D Usha Nandini
Sathyabama University, Chennai, Tamil Nadu, IN

M Saravanan
Department of Computer Science and Engineering, Sathyabama University, Chennai

J Albert Mayan
Sathyabama University, Chennai, Tamil Nadu, IN

Murari Devakannan Kamalesh
Sathyabama University, Chennai, Tamil Nadu, IN

K Mohana Prasad
Sathyabama University, Chennai, Tamil Nadu, IN

Automatic toll payment, alcohol detection, load and vehicle information using Internet of things & mailing system

Publisher: IEEE Cite This PDF

Anand Kumar ; N Anusha ; B Sai Srinivasa Vara Prasad All Authors

Abstract:
Present days, the technology has emerged in such a way that it can provide solutions to any real life problems. One such solution to the problem of collecting toll payments efficiently is RFID. When a vehicle passes through the toll gate, by using RFID the vehicle can be detected easily and the amount will be deducted from the user's prepaid account automatically. In order to detect the drunken drivers, Alcohol sensor (MQ135) is fixed near the driver steering. If the alcohol consumed by the driver exceeds the limit detected by sensors automatically accelerometer reduces the speed. The load detector fixed at the toll gates checks the load of heavy vehicles if load exceeds the limit, the barrier remains closed and the camera captures a snap of the vehicle. And also the notification will be sent to the consignment officer for registering complaint against the vehicle and also penalty amount will be deducted from the owner's account.

Published in: 2017 International Conference on Intelligent Computing and Control (I2C2)

Date of Conference: 23-24 June 2017

INSPEC Accession Number: 17634442

Date Added to IEEE *Xplore*: 22 March 2018

DOI: 10.1109/I2C2.2017.8321775

ISBN Information:

Electronic ISBN:978-1-5386-0374-1

Print on Demand(PoD) ISBN:978-1-5386-0375-8

Publisher: IEEE

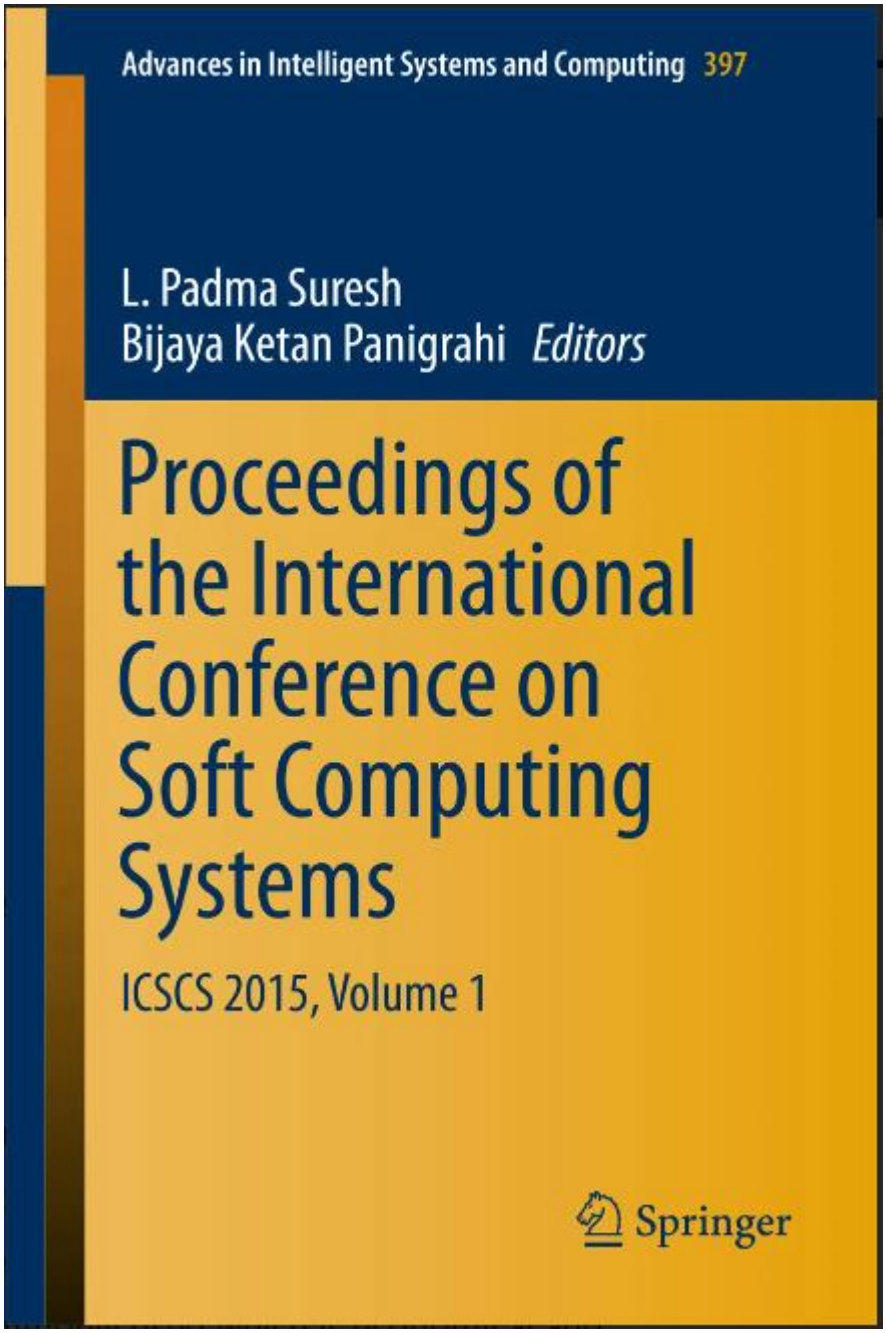
Conference Location: Coimbatore, India

Authors

Anand Kumar
Computer Science Engineering Sathyabama University, Chennai

N Anusha
Computer Science Engineering Sathyabama University, Chennai

B Sai Srinivasa Vara Prasad
Computer Science Engineering Sathyabama University, Chennai



Application of DSTATCOM for Loss Minimization in Radial Distribution System

P. Shanmugasundaram ✉ & A. Ramesh Babu

Conference paper | First Online: 29 December 2015

1676 Accesses | **4** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 397)

Abstract

Power system consists of three parts, first one is generation, second one is transmission, and third one is distribution. The distribution system provides electric power to end consumer. The distribution line has three configurations. They are radial, two end fed, and ring distribution system. Among them radial distribution is most commonly used which is economical and rugged system. This radial distribution system is suffers by loss. This loss can be reduced by including capacitor in the system which is discrete size and not suitable for fine adjustment and it is a passive device and has slow response. Distributed static compensator (DSTATCOM) is flexible AC transmission system (FACTS) device which is fast and continuous adjustable device. The DSTATCOM provide adequate reactive power support and reduce losses in the radial distribution system. To demonstrate the effect of DSTATCOMIEEE 33 Bus test system is considered.

Keywords

DSTATCOM **Loss minimization** **VAR injection** **FACT** **Sensitive bus**

Author information

Authors and Affiliations

Department of EEE, Sathyabama University, Chennai, India

P. Shanmugasundaram & A. Ramesh Babu

Corresponding author

Correspondence to P. Shanmugasundaram .

Editor information

Editors and Affiliations

Noorul Islam Centre for Higher Education, Kumaracoil, Tamil Nadu, India

L. Padma Suresh

IIT Delhi, New Delhi, Delhi, India

Bijaya Ketan Panigrahi

Cite this paper

Shanmugasundaram, P., Ramesh Babu, A. (2016). Application of DSTATCOM for Loss Minimization in Radial Distribution System. In: Suresh, L., Panigrahi, B. (eds) Proceedings of the International Conference on Soft Computing Systems. Advances in Intelligent Systems and Computing, vol 397. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2671-0_18

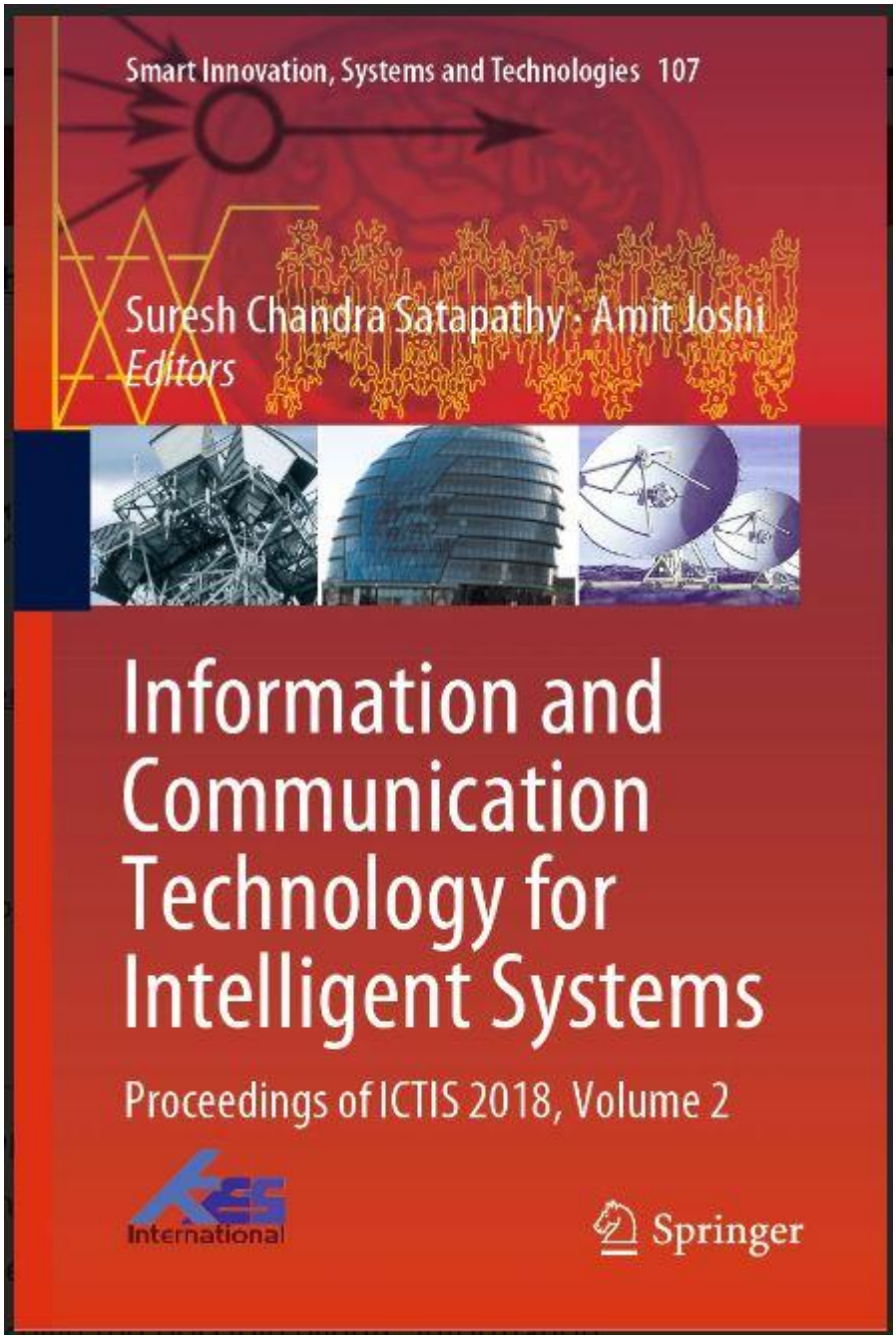
Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-81-322-2671-0_18

Published	Publisher Name	Print ISBN
29 December 2015	Springer, New Delhi	978-81-322-2669-7
Online ISBN	eBook Packages	
978-81-322-2671-0	Engineering Engineering (R0)	


Smart Innovation, Systems and Technologies 107
Suresh Chandra Satapathy · Amit Joshi
Editors
Information and Communication Technology for Intelligent Systems
Proceedings of ICTIS 2018, Volume 2
KES International
Springer



Information and Communication Technology for Intelligent Systems pp 675–682 | Cite as

Anomaly Detection by Using CFS Subset and Neural Network with WEKA Tools

J. Jabez ✉, S. Gowri, S. Vigneshwari, J. Albert Mayan & Senduru Srinivasulu

Conference paper | First Online: 15 December 2018

1229 Accesses | **15** Citations

Part of the Smart Innovation, Systems and Technologies book series (SIST,volume 107)

Abstract

An intrusion detection system (IDS) is a product application or contraption that screens the framework or system practices for methodology encroachment or dangerous activities and makes reports to the organizational framework. The principal centralization of intrusion discovery and aversion frameworks (IDPS) is to perceive the possible events, information logging about them, and interruption tries to report. Furthermore, the associations are utilizing IDPS for different purposes, such as recognizing issues identified with approaches of security, recording, and keeping the people existing dangers from encroaching arrangements of security. In this paper, anomaly is identified utilizing enhanced correlation-based feature selection (CFS), which is basically a subset technique and is based upon extreme learning machine, multilayer perceptron, and feature selection. This project scope involves identification of anomalies in the early stages and to increase the accuracy of identification or detection.

Keywords

IDS **Feature selection** **Anomaly** **Multilayer perceptron** **EML**

Authors and Affiliations

School of Computing, Sathyabama Institute of Science & Technology, Chennai, Tamil Nadu, India

J. Jabez, S. Gowri, S. Vigneshwari, J. Albert Mayan & Senduru Srinivasulu

Corresponding author

Correspondence to J. Jabez .

Editor information

Editors and Affiliations

School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India

Dr. Suresh Chandra Satapathy

Sabar Institute of Technology, Gujarat Technological University, Ahmedabad, Gujarat, India

Dr. Amit Joshi

Cite this paper

Jabez, J., Gowri, S., Vigneshwari, S., Albert Mayan, J., Srinivasulu, S. (2019). Anomaly Detection by Using CFS Subset and Neural Network with WEKA Tools. In: Satapathy, S., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems . Smart Innovation, Systems and Technologies, vol 107. Springer, Singapore. https://doi.org/10.1007/978-981-13-1747-7_66

Download citation

.RIS .ENW .BIB

DOI

https://doi.org/10.1007/978-981-13-1747-7_66

Published	Publisher Name	Print ISBN
15 December 2018	Springer, Singapore	978-981-13-1746-0
Online ISBN	eBook Packages	
978-981-13-1747-7	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)	

Animal tracking using background subtraction on multi threshold segmentation

Publisher: IEEE Cite This PDF

E. Surendar ; Vimin M Thomas ; A. Mary Posonia **All Authors**

Abstract:
The proposed project performs motion detection and animal tracking based on background subtraction using multi threshold approach with mathematical morphology. Here the techniques frame differences; multi threshold based detection will be used. Along with this multi threshold, mathematical morphology also used which has an ability to attenuate color variations produced by background motions which will highlight moving objects. After the object foreground detection, the parameters like animal or human will be detected. Finally the simulated results will be shown. It will use approximate method with mathematical morphology approach rather than prior background subtraction methods using animal tracking on efficient result.

Published in: 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)

Date of Conference: 18-19 March 2016

Date Added to IEEE *Xplore*: 04 August 2016

ISBN Information:

Electronic ISBN:978-1-5090-1277-0

DVD ISBN:978-1-5090-1276-3

Print on Demand(PoD) ISBN:978-1-5090-1278-7

INSPEC Accession Number: 16195331

DOI: 10.1109/ICCPCT.2016.7530223

Publisher: IEEE

Conference Location: Nagercoil, India

Authors

E. Surendar
Department of Computer Science and Engineering, Sathyabama University, Chennai, India

Vimin M Thomas
Department of Computer Science and Engineering, Sathyabama University, Chennai, India

Analysis of time bound collaborative access control delegation model in electronic health records

Publisher: IEEE Cite This PDF

C. Kalpana ; S. Revathy All Authors

Abstract:
Electronic Health Records (EHR) is the digital record shared across multiple users like hospital, personal physicians, researchers, insurance companies and many. Such a scenario of data sharing attracts much attention towards security and privacy of their own medical records. These data are patient centric and patients are very much concerned towards the limitations in sharing of their own medical records. We present a time bound combined access control model which provides concerns on spatial dimensions and time constraints to concern over temporal dimensions. This model aims in providing privacy since dynamic key creation and distribution is included based on the time constraint using a new key assignment scheme suitable for hierarchical systems. The research focus will be on insisting the importance of the role of Delegation and Access Control in Health care. Time Bound Access Control is to be proposed and concentrated much on developing and implementing algorithms related to combining various access controls techniques as a time bound delegation model. Combining privacy policy through an analyzer will control the access of data, which gives the privacy of the data.

Published in: 2017 International Conference on Computing Methodologies and Communication (ICCMC)

Date of Conference: 18-19 July 2017

INSPEC Accession Number: 17575209

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCMC.2017.8282724

ISBN Information:

Electronic ISBN:978-1-5090-4890-8

Print on Demand(PoD) ISBN:978-1-5090-4891-5

Publisher: IEEE

Conference Location: Erode, India

Authors

C. Kalpana
Department of Computer Applications, Rajiv Gandhi College of Engineering and Technology, Puducherry, India

S. Revathy
Department of Information Technology, Sathyabama University, Chennai, India

An Embedding of Secret Message in Audio Signal

Publisher: IEEE

Cite This PDF

Alaknanda S. Patil ; G. Sundari **All Authors**

Abstract:
The steganography is the embedding process of the secret message in the other carrier message. It carries the security, transparency and authenticity. The type of steganography is depends on the cover message used, if cover message used is audio to hide any type of data, it is audio steganography. The secret data hidden behind the audio can be image, audio or text. Audio steganography is more preferable because of the lots of redundancy. The different domains and methodologies are in exist for secret data embedding in audio signal. This paper gives the variety of algorithms to study the Audio Steganography.

Published in: 2018 3rd International Conference for Convergence in Technology (I2CT)

Date of Conference: 06-08 April 2018

INSPEC Accession Number: 18249178

Date Added to IEEE *Xplore*: 11 November 2018

DOI: 10.1109/I2CT.2018.8529549

ISBN Information:

Electronic ISBN:978-1-5386-4273-3

CD:978-1-5386-4272-6

Print on Demand(PoD) ISBN:978-1-5386-4274-0

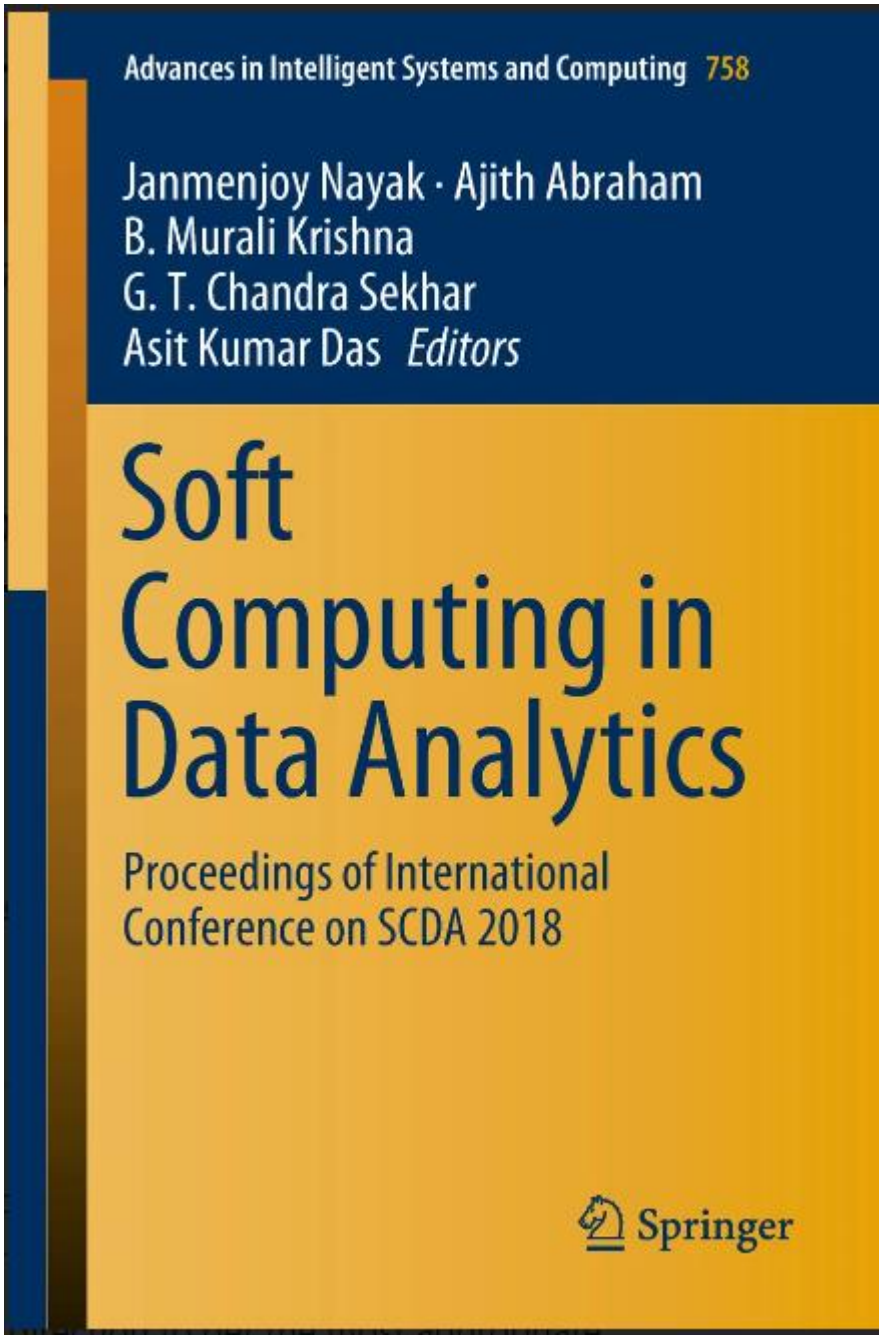
Publisher: IEEE

Conference Location: Pune, India

Authors

Alaknanda S. Patil
Electronics Engineering Department, Sathyabama University, Chennai, TN, India

G. Sundari
Electronics Engineering Department, Sathyabama University, Chennai, TN, India


Advances in Intelligent Systems and Computing 758
Janmenjoy Nayak · Ajith Abraham
B. Murali Krishna
G. T. Chandra Sekhar
Asit Kumar Das Editors
Soft Computing in Data Analytics
Proceedings of International Conference on SCDA 2018
Springer



Soft Computing in Data Analytics pp 735–742 | Cite as

An Efficient Framework for Document Retrieval in Relationship Strengthened Personalized Ontologies

S. Vigneshwari, A. Mary Posonia & S. Gowri

Conference paper | First Online: 22 August 2018

684 Accesses | **1** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 758)

Abstract

Personalization is crucial in the prevailing Internet scenario. An efficient information retrieval not only relies on personalization but also upon security. There is need of a semantic approach of document retrieval to boost the search quality. The incorporation of knowledge extracted from multiple ontologies is the significant research direction to get the most appropriate documents from the knowledge base. In this paper, local and global ontologies are used for mining the semantic associations. Multiple ontologies can share a variety of concepts, and automatic merging of them is feasible. Any input query is often matched with the semantic association for finding the relevant words. This is achieved by analyzing the intensified concept relationships in the local repository or local ontology and finding out the possibilities of automation by similarity calculation. A user profiling model is constructed which is trained for personalization. A novel framework has been proposed which provides document retrieval. The results are analyzed using benchmark SWETO ontology. Due to the query property enhancements, the error rate is reduced and the local ontology gets optimized for synchronization with the global Web ontology, thereby enhancing user personalization.

Keywords

Local ontology **Global Web ontology** **Query enhancement** **Automated ontologies** **Document retrieval** **Security** **Cloud**

Authors and Affiliations

School of Computing, Sathyabama Institute of Science & Technology, Chennai, 600119, Tamil Nadu, India

S. Vigneshwari, A. Mary Posonia & S. Gowri

Corresponding author

Correspondence to S. Vigneshwari .

Editor information

Editors and Affiliations

Department of Computer Science and Engineering, Sri Sivani College of Engineering, Srikakulam, Andhra Pradesh, India

Dr. Janmenjoy Nayak

Machine Intelligence Research Labs (MIR Labs), Scientific Network for Innovation and Research Excellence, Washington, USA

Prof. Dr. Ajith Abraham

Department of Mechanical Engineering, Sri Sivani College of Engineering, Srikakulam, Andhra Pradesh, India

Dr. B. Murali Krishna

Department of Electrical and Electronics Engineering, Sri Sivani College of Engineering, Srikakulam, Andhra Pradesh, India

Dr. G. T. Chandra Sekhar

Department of Computer Science and Technology, Indian Institute of Engineering Science and Technology (IIEST), Shibpur, Howrah, West Bengal, India

Dr. Asit Kumar Das

Cite this paper

Vigneshwari, S., Mary Posonia, A., Gowri, S. (2019). An Efficient Framework for Document Retrieval in Relationship Strengthened Personalized Ontologies. In: Nayak, J., Abraham, A., Krishna, B., Chandra Sekhar, G., Das, A. (eds) Soft Computing in Data Analytics . Advances in Intelligent Systems and Computing, vol 758. Springer, Singapore.
https://doi.org/10.1007/978-981-13-0514-6_70

Download citation

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-13-0514-6_70

Published	Publisher Name	Print ISBN
22 August 2018	Springer, Singapore	978-981-13-0513-9
Online ISBN	eBook Packages	
978-981-13-0514-6	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)	

IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

An Advanced Attendance Marking system using facial Recognition

Jarugula Vamsikrishna[1], Kollipara Anudeep[1], L. Jegan Antony Marcilin[2] and V. Balamurugan[3]

Published under licence by IOP Publishing Ltd

IOP Conference Series: Materials Science and Engineering, Volume 590, International Conference on Frontiers in Materials and Smart System Technologies 10 April 2019, Tamil Nadu, India

Citation Jarugula Vamsikrishna *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **590** 012049

DOI 10.1088/1757-899X/590/1/012049

Abstract

With advancement in computing and telecommunications technologies, digital images and videos are playing main role in the current information era. Human face is an important bio-metric object in image and video database of surveillance systems. Detecting and location of human faces and facial features in an image sequence is an important task in dynamic environments, such as videos where noise conditions, illuminations, locations of subjects and pose can vary significantly from frame to frame. An automated attendance system is used for marking human face recognition in real time for college to mark the attendance for their staff and students. Smart automated attendance marking using real time face recognition is a real world solution which comes with day to day activities of handling employees. Here multiple user faces are detected and recognized with the data base trained multiple texture based notability.

IOP Conf. Series: Materials Science and Engineering **590** (2019) 012049 doi:10.1088/1757-899X/590/1/012049

An Advanced Attendance Marking system using facial Recognition

Jarugula.Vamsikrishna[1]
Final Year UG student-Dept of ECE,
Sathyabama Institute of Science and Technology,
Chennai – India.
vamsikrishna0777@gmail.com

Kollipara.Anudeep[1]
Final Year UG student-Dept of ECE,
Sathyabama Institute of Science and Technology,
Chennai – India.
kolliparaanudeep1234@gmail.com

L.Jegan Antony Marcilin[2]
Assistant Professor, Department of ECE
Sathyabama Institute of Science and Technology,
Chennai – India.
jegan25@hotmail.com

V. Balamurugan[3]
Assistant Professor, Department of ECE
Sathyabama Institute of Science and Technology,
Chennai – India.
vbbala@gmail.com.

A Women Safety Portable Hidden camera detector and jammer

Publisher: IEEE Cite This PDF

Sindhu K ; R. Subhashini ; S. Gowri ; J. S. Vimali All Authors

Abstract:
In recent times our country had developed in many ways but still many crimes against women exist. There are no proper means to detect the hidden camera which now had become a serious issue. This device intends to maintain the safety and security of people by detecting the spy camera and stop the camera working. Since hidden cameras are less cost and small in size, it became easy to install camera almost everywhere. It finds its application in the areas where actually cameras are not allowed. This device works with a principle that camera operates in the frequency range 0.1GHz to 3GHz, so RF signal detector detects the frequency range and if it is within this frequency range it alerts us with an alert light. A video camera can also be detected by this detector by extending the detecting range from 0.05 GHz to 3 GHz we can detect both camera and video camera. Once, it is detected it alerts us with a LED light and it an option to stop the camera transmission by creating strong interference between the camera and base station. This device also comes along with an eyeglass which can be used if alternating pulsating frequency or if recording camera is used where RF signal detector cannot be used. Cameras can be detected by flashing red light once viewed through optical glass viewport. The flashing light will reflect off the `spy' camera's lens and will be able to spot it.

Published in: 2018 3rd International Conference on Communication and Electronics Systems (ICCES)

Date of Conference: 15-16 October 2018

Date Added to IEEE *Xplore*: 30 May 2019

ISBN Information:
Electronic ISBN:978-1-5386-4765-3
Print on Demand(PoD) ISBN:978-1-5386-4766-0

INSPEC Accession Number: 18724013

DOI: 10.1109/CESYS.2018.8724066

Publisher: IEEE

Conference Location: Coimbatore, India

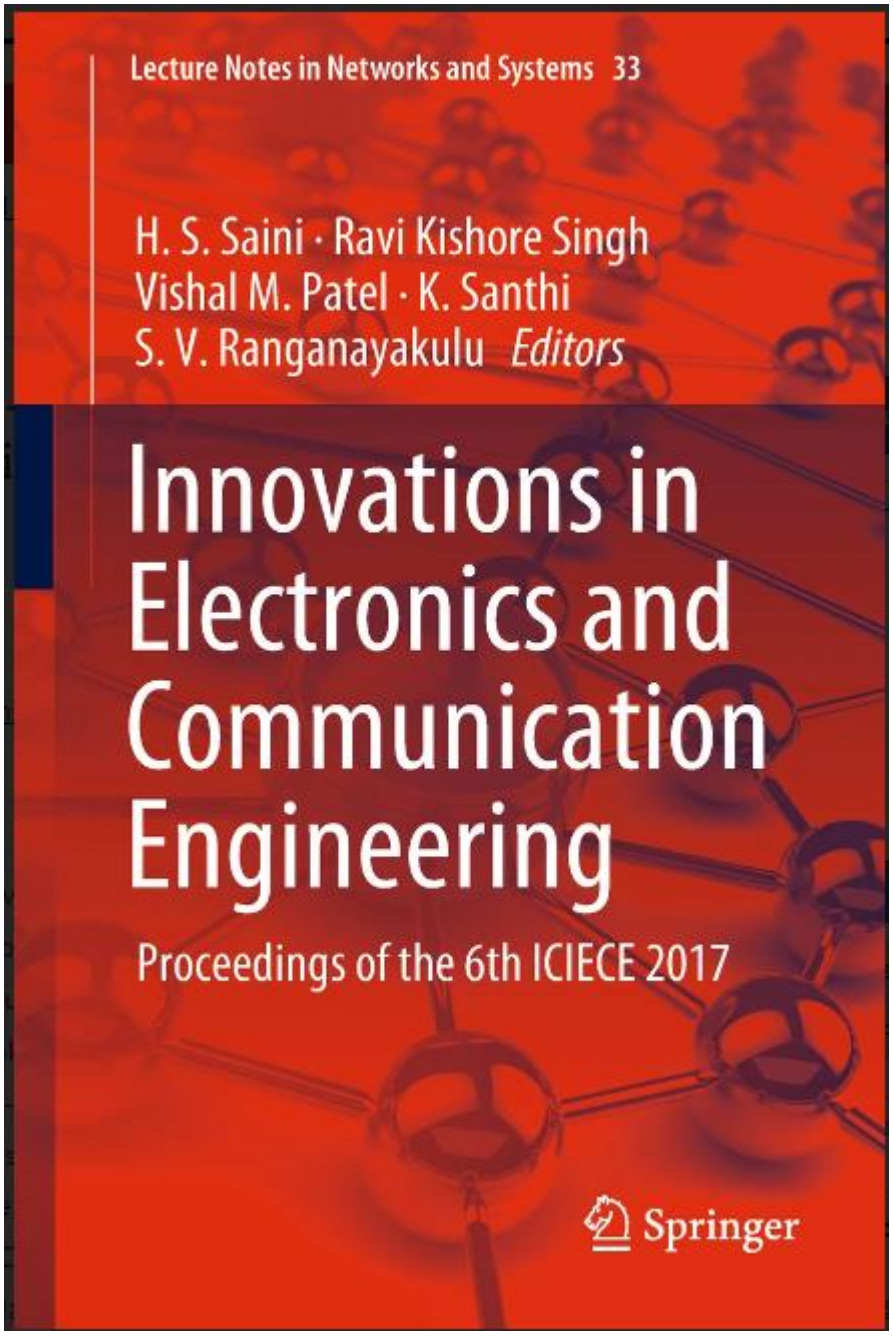
Authors

Sindhu K
Department of IT, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

R. Subhashini
Department of IT, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

S. Gowri
Department of IT, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

J. S. Vimali
Department of IT, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India


Lecture Notes in Networks and Systems 33
H. S. Saini · Ravi Kishore Singh
Vishal M. Patel · K. Santhi
S. V. Ranganayakulu Editors
Innovations in Electronics and Communication Engineering
Proceedings of the 6th ICIECE 2017
Springer



Innovations in Electronics and Communication Engineering pp 159–165 | Cite as

A Virtual Instrumentation-Based Data Acquisition System for Vibration Monitoring and Analysis

S. M. Shyni, K. Vasanth & C. Bhuvaneswari

Conference paper | First Online: 29 August 2018

539 Accesses

Part of the Lecture Notes in Networks and Systems book series (LNNS,volume 33)

Abstract

The proposed work aims at monitoring the vibration of the bridge in virtual instrumentation namely LabVIEW for finding the age of the bridge ensuring the safety. The vibration is measured by the vibration sensor and the output is monitored through the PC by virtual instrumentation method. Embedded Controller is programmed to read the sensor value. The vibration sensor is connected to the Analog-to-Digital Converter of the controller through the signal conditioning. Then, the output is transmitted to the PC through the wireless technology. The firmware for the microcontroller is done by embedded 'C' program. The LabVIEW software processes the data from serial port of the PC and takes the appropriate precautions and actions to prevent the bridge from further damage.

Keywords

Data acquisition | Monitoring | Vibration | Sensors | LabVIEW

Authors and Affiliations

Department of Electrical and Electronics Engineering, Sathyabama University, Chennai, India

S. M. Shyni & C. Bhuvaneswari

Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

K. Vasanth

Corresponding author

Correspondence to S. M. Shyni .

Editor information

Editors and Affiliations

Guru Nanak Institutions, Ibrahimpatnam, Telangana, India

Dr. H. S. Saini

Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

Dr. Ravi Kishore Singh

Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, NJ, USA

Prof. Vishal M. Patel

Department of Electronics and Communication Engineering, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

Dr. K. Santhi

Research and Development, Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India

Prof. S.V. Ranganayakulu

Cite this paper

Shyni, S.M., Vasanth, K., Bhuvaneswari, C. (2019). A Virtual Instrumentation-Based Data Acquisition System for Vibration Monitoring and Analysis. In: Saini, H., Singh, R., Patel, V., Santhi, K., Ranganayakulu, S. (eds) Innovations in Electronics and Communication Engineering. Lecture Notes in Networks and Systems, vol 33. Springer, Singapore. https://doi.org/10.1007/978-981-10-8204-7_16

Download citation

.RIS .ENW .BIB

DOI
https://doi.org/10.1007/978-981-10-8204-7_16

Published	Publisher Name	Print ISBN
29 August 2018	Springer, Singapore	978-981-10-8203-0
Online ISBN	eBook Packages	
978-981-10-8204-7	Engineering Engineering (R0)	

A Survey: Integration of IoT and Fog Computing

Publisher: IEEE

Cite This

PDF

M. Jalasri ; L. Lakshmanan **All Authors**

Abstract:
Internet of Things (IoT) and Fog Computing (FC) are rapid development in the field of wireless network. Features such as decentralization, low latency, and location awareness are not supported by IoT which is significant for some IoT applications, including smart shopping cart, smart meter, and smart home. These features are supported by the amalgamation of IoT and FC to extend services such as storage, computing, and networking resources to the network edge. This paper presents a survey of IoT and FC with security threats and also the benefits of their combination. Finally, the security challenges of IoT and FC have studied.

Published in: 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT)

Date of Conference: 16-18 August 2018

Date Added to IEEE *Xplore*: 08 July 2019

ISBN Information:
Electronic ISBN:978-1-5386-5657-0
USB ISBN:978-1-5386-5656-3
Print on Demand(PoD) ISBN:978-1-5386-5658-7

INSPEC Accession Number: 18797442

DOI: 10.1109/ICGCIoT.2018.8753010

Publisher: IEEE

Conference Location: Bangalore, India

Authors

M. Jalasri
Department of CSE, Sathyabama Institute of Science And Technology, Chennai, India

L. Lakshmanan
Department of CSE, Sathyabama Institute of Science And Technology, Chennai, India

Received January 30, 2021, accepted February 27, 2021, date of publication March 2, 2021, date of current version March 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3063615

Implementing Big Data Analytics in E-Commerce: Vendor and Customer View

SARAH S. ALRUMIAH[1] AND MOHAMMED HADWAN[1,2,3], (Member, IEEE)

[1]Department of Information Technology, College of Computer, Qassim University, Buraydah 51452, Saudi Arabia
[2]Intelligent Analytics Group (IAG), College of Computer, Qassim University, Buraydah 51452, Saudi Arabia
[3]Department of Computer Science, College of Applied Sciences, Taiz University, Taiz, Yemen

Corresponding author: Sarah S. Alrumiah (saraalrumih@gmail.com)

ABSTRACT Due to the information revolution, one of the biggest challenges of e-commerce is the huge data that needs to be processed and analyzed to gain benefits from it. Big Data Analytics (BDA) aims to improve the decision-making process by analyzing and understanding big data, e.g., messages, social media posts, etc. Furthermore, BDA capabilities are used in e-commerce activities as a key growth direction to increase vendors' revenues and attract customers. Based on the importance of analyzing big data and its advantages to e-commerce operations, this paper aims to study the values of implementing BDA in e-commerce to both vendors and consumers. Fifteen papers are selected to analyze the impacts of analyzing big data in e-commerce. Electronic vendors (E-vendors) use BDA to gain the competitive advantages they need to understand consumer behaviour and increase their income by improving customer loyalty. Besides, recommendation systems derived from BDA personalize the searching and shopping experience of the customers. However, there are some negative effects derived from applying BDA in e-commerce, such as shopping addiction. In addition, e-vendors have to deal with expensive BDA tools and professionals. In conclusion, even though BDA enhances consumers' and vendors' electronic shopping experience, the rapid growth of data is still challenging.

INDEX TERMS Big data analytics (BDA), E-commerce, predictive analytics, customization, increase revenues.

Received January 30, 2021, accepted February 27, 2021, date of publication March 2, 2021, date of current version March 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3063615

A survey on rise of mobile malware and detection methods

Publisher: IEEE Cite This PDF

S. Kalpana ; S. Karthikeyan All Authors

Abstract:
Mobile phones are indispensable in our day to day live as it allows us to access enormous services. Hence in recent days the mobile devices are otherwise called as smart phones. The wireless communication technology ranges from GSM, GPRS, Bluetooth, Wi-Fi, 3G and the latest is 4G. It offers attractive features which allow internet access, online banking, install any numbers of Apps. These features attract the mobile user community to shift from traditional phone to smart phone and the Malware writers too. In this paper, we have outlined the evolution and wide spread of mobile malwares, attack vectors, various methods to detect malwares.

Published in: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

Date of Conference: 17-18 March 2017

Date Added to IEEE *Xplore*: 01 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-3294-5
Print on Demand(PoD) ISBN:978-1-5090-3295-2

INSPEC Accession Number: 17558836

DOI: 10.1109/ICIIECS.2017.8276158

Publisher: IEEE

Conference Location: Coimbatore, India

Authors

S. Kalpana
Department of Information Technology, Sathyabama University, Chennai, India

S. Karthikeyan
Department of Information Technology, Sathyabama University, Chennai, India

A Survey on IoT Technologies, Evolution and Architecture

Publisher: IEEE

Cite This PDF

R.M. Gomathi ; G. Hari Satya Krishna ; E. Brumancia ; Y. Mistica Dhas **All Authors**

Abstract:
The Internet Of Things (IoT) was initially called as “Internet of Everything”. IoT objects have identifiers which are unique. They have the ability to send the information over a network interconnecting physical and virtual things without requiring human-to-human or human-to-system interaction. It provides the effective way of learning and interaction among internet connected devices that include sensors, actuators, services and other internet connected objects. The IoT can be defined as things or objects connected to the internet than people. The IoT system's middleware can be described as intermediary software system between the IoT devices and applications. The IoT aim is to unite everything in the world under common infrastructure. The main objective of this paper is to provide an survey on Internet of Things, types of architectures, and the technologies that are used in our daily life, smart environment. However, this paper gives a good idea for new researches on Internet of Things.

Published in: 2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP)

Date of Conference: 22-23 February 2018

Date Added to IEEE *Xplore*: 02 September 2018

ISBN Information:
Electronic ISBN:978-1-5386-1141-8
USB ISBN:978-1-5386-1140-1
Print on Demand(PoD) ISBN:978-1-5386-1142-5

INSPEC Accession Number: 18075535

DOI: 10.1109/ICCCSP.2018.8452820

Publisher: IEEE

Conference Location: Chennai, India

Authors

R.M. Gomathi
Dept. of CSE, Sathyabama University, Chennai, India

G. Hari Satya Krishna
Dept. of IT, Sathyabama University, Chennai, India

E. Brumancia
Dept. of CSE, Sathyabama University, Chennai, India

Y. Mistica Dhas
Dept. of CSE, Sathyabama University, Chennai, India

A study on denoising techniques eliminating impulse noise and artifacts

Publisher: IEEE Cite This PDF

J. Jenifer ; K. Vasanth ; G. Harshitha ; N. M. Nandhitha All Authors

Abstract:
Image is a potential standard to convey information. Impulse noise degrades the image during image transmission and acquisition. The most important task of image processing is noise filtering and image enhancement. De-noising the degraded image is an important research region in Image Processing. Many de-noising proposals established which used standard median filter or its alterations to suppress noise. In this analysis paper, different algorithms for exclusion of different impulse noise and their performance over different noise circumstances are discussed. Each and every scheme has its benefits and drawbacks. The relative study assist in detection of most proficient algorithms in terms of edge preservation and noise suppression to restore the original image to the finest potential level. Most of the systems are suitable for some particular noise models, and does not perform effectively on other kind of noise models.

Published in: 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)

Date of Conference: 23-24 March 2017

Date Added to IEEE *Xplore*: 18 January 2018

ISBN Information:
Electronic ISBN:978-1-5090-4855-7
Print on Demand(PoD) ISBN:978-1-5090-4856-4

INSPEC Accession Number: 17521202

DOI: 10.1109/ICONSTEM.2017.8261345

Publisher: IEEE

Conference Location: Chennai, India

Authors

J. Jenifer
Electronics and Communication Engineering, Sathyabama University, Chennai, Tamil Nadu, India

K. Vasanth
Electronics and Communication Engineering, Vidhya Jyothi Institute of Technology, Telangana, Hyderabad, India

G. Harshitha
Electronics and Communication Engineering, Sathyabama University, Chennai, Tamil Nadu, India

N. M. Nandhitha
Electronics and Communication Engineering, Sathyabama University, Chennai, Tamil Nadu, India

Lecture Notes in Mechanical Engineering

Puneet Verma
Olusegun D. Samuel
Tikendra Nath Verma
Gaurav Dwivedi *Editors*

Advancement in Materials, Manufacturing and Energy Engineering, Vol. II

Select Proceedings of ICAMME 2021

Springer

Editors
Puneet Verma
School of Earth and Atmospheric Science
Queensland University of Technology
Brisbane, QLD, Australia

Olusegun D. Samuel
Department of Mechanical Engineering
Federal University of Petroleum Resource
Effurun, Nigeria

Tikendra Nath Verma
Department of Mechanical Engineering
Maulana Azad National Institute
of Technology
Bhopal, India

Gaurav Dwivedi
Energy Centre
Maulana Azad National Institute
of Technology
Bhopal, India

ISSN 2195-4356 ISSN 2195-4364 (electronic)
Lecture Notes in Mechanical Engineering
ISBN 978-981-16-8340-4 ISBN 978-981-16-8341-1 (eBook)
https://doi.org/10.1007/978-981-16-8341-1

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2022
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar

Air Quality Assessment During Lockdown for Jaipur City 107
Darshan Mishra, Kaushiki Sharma, and Sohil Sisodiya

Micellar Potentiometric Determination of Stability Constant and Antibacterial Investigations on Complexes of Cu (II) with 3, 4, 5-Trimethoxybenzaldehyde Thiosemicarbazone 123
Uma Rathore, Raja Ram, Kamal Kishor Verma, and N. Bhojak

Aging of a Thermoplastic Polymer Under the Combined Effect of Ultraviolet Ray and Temperature 135
Sonya Redjala, Said Azem, and Nourredine Ait Hocine

Effects of Additives and Treatment on Fly Ash-Based Polymer Composites 147
Sushant Patel, G. L. Devnani, and Deepesh Singh

Powder Mixed Electrical Discharge Machining of EN 31 Steel 155
Rakesh Kumar Patel and M. K. Pradhan

Simulation and Analysis of Hybrid Energy Resources in Chennai Using Homer Software 163
P. Abirami, M. Pushpavalli, V. Geetha, P. Sivagami, and R. Harikrishnan

EDM Process Optimization of Machining Parameters for Through Hole Making on HCHCR AISI-D7 Steel Using RSM 179
Amit Kumar and Mohan K. Pradhan

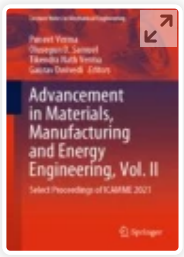
Improving the Performance of Supply Chain Through Industry 4.0 Technologies 197
Mohd Ammar, Abid Haleem, Mohd Javaid, Shashi Bahl, and Devaki Nandan

Significant Applications of Composite and Natural Materials for Vibration and Noise Control: A Review 211
Kartikay Singh Pawar, Ashok Kumar Bagha, Shashi Bahl, and Devaki Nandan

Digital Management Systems in Manufacturing Using Industry 5.0 Technologies 221
Nadia Fazal, Abid Haleem, Shashi Bahl, Mohd Javaid, and Devaki Nandan

Springer Link

Search Log in



Advancement in Materials, Manufacturing and Energy Engineering, Vol. II pp 163–177

Simulation and Analysis of Hybrid Energy Resources in Chennai Using Homer Software

P. Abirami, M. Pushpavalli, V. Geetha, P. Sivagami & R. Harikrishnan

Conference paper | First Online: 18 January 2022

507 Accesses

Part of the Lecture Notes in Mechanical Engineering book series (LNME)

Abstract

The vision of our nation is to implement renewable energy resources in a wide range for all industrial and commercial processes. The efficient way of extracting power from renewable sources is obtained by interconnecting two or more resources. Emerging renewable energy sources in power sector are wind and solar since they are inexhaustible sources. These resources will help in attaining future energy demands. So, in this research paper, solar and wind sources are interconnected through buses. Wind source is connected with AC bus and Solar is connected with

Health Information Science



Amit Kumar Manocha
Shruti Jain
Mandeep Singh
Sudip Paul *Editors*

Computational Intelligence in Healthcare

Springer

4th International Conference on Trends in Electronics and Informatics

(ICOEI 2020)

Table of Contents

S. No	Paper Title/ Author Name	Page. No
1	Test Scheduling for Low Transition Reusable LFSR based BIST in 3-D Stacked ICs Bandi Divya, Navya Mohan	1
2	Optical Signal Confinement in an Optical Sensor for Efficient Biological Analysis by HQF Achievement Sowmya Padukone G, Uma Devi H	7
3	Stimulating Neurons with Another Activation Function Darshak B. Mehta, Pratik A. Barot, Shaily G. Langhnoja	13
4	8-Bit ALU Design Using m-GDI Technique Shubham Sarkar, Hijal Chatterjee, Pritam Saha, Manoj Biswas	17
5	User Request Emotion Prediction Approach in A Crowdsourcing Platform Priyanga K M, Alamelu M	23
	Navigation by Landmark Recognition Using Camera 360°	

162	Anisotropic Smoothing for Illumination Invariant Face Anti-spoofing Raghavendra R J, Sanjeev Kunte R	909
163	Improving Ranking in Document based Search Systems Remya R.K. Menon, Jagan Kaartik, Karthik Nambiar E T, Athul Krishna T K, Arun Kumar S	914
164	Classification of Cardiac Arrhythmia using Kernelized SVM Yogita Bhatia, Akanksha Mittal, Shefali Athavale, Tanya Mohanani, Gresha Bhatia	922
165	Machine Learning based RecyClick: Recycle At A Click Kaustubh Acharekar, Pranav Khedekar, Joanna Dsouza, Sonali Vaidya	927
166	Reducing Overfitting Problem in Machine Learning Using Novel L1/4 Regularization Method Johnson Kolluri, Vinay Kumar Kotte, Phridviraj M S B, Shaik Razia	934
167	Applying Halstead Software Science on Different Programming Languages for Analyzing Software Complexity Nikhil Govil	939
168	Automation of Blood Donation by Data Integration using Data Mining Murari Devakannan Kamalesh, Albert Mayan J, Yovan Felix, Dhamodaran S, Mohana Prasad	944
169	Review of Machine Learning Methods for the Survey on HCC Scenario and Prediction Strategy Vyshali J Gogi, Vijayalakshmi M N	949
170	On Trend Hybrid Pattern Generation in Fabrics using Machine Learning Sandhya Viswanath, Rajathi N	952
171	Analysis of Intelligent Machines using Deep learning and Natural Language Processing Manjunath R Kounte, Pratyush Kumar Tripathy, Pramod P, Harshit Bajpai	956
172	Development of Non Invasive Cardiovascular Risk Diagnosing System Kalyan K Reddy, Abilash K V, Sridevi S, Pappa M, Sridhar N, Sateesh A	961

Automatic Melanoma Diagnosis and Classification on Dermoscopic Images



Bethanney Janney. J, S. Emalda Roslin, and J. Premkumar

Abstract Automatic identification of skin carcinoma is significant. The dangerous types of cancer found in humans are skin cancer. Melanoma is the deadlier type of skin cancer. Healing can be beneficial on early diagnosis of melanoma. This paper aims to propose a skin cancer diagnosis algorithm that can automatically classify lesions as malignant or benign. Segmentation is done in the initial phase after choosing the finest image development methods accomplished by applying and comparing various noise removal filters on images. A perfect optimized segmentation methodology relies on thresholding, watershed algorithms are implemented on dermoscopy images, and the differentiated object is derived for GLCM attributes and qualitative techniques. To identify malignant or benign samples, the attributes are therefore supplied in to the support vector machine as a source. As perceived by experimental findings, SVM classification system with watershed segmentation has the potential to distinguish benign and malignant melanoma with a higher overall categorization efficiency.

Keywords Dermoscopic skin lesions · Image enhancement · Watershed segmentation · Feature extraction · SVM classifier

1 Introduction

Of most of the recognized forms of cancer, skin cancer, is considered as well established. It is frequently seen in the Caucasian white people. Melanoma, basal skin cancer (BSC), and squamous skin cancer (SSC) are predominant forms of skin cancer. Between these three, both basal and squamous are termed non-melanocytic skin cancer (NMSC) unlike the other skin conditions [11]. Melanoma seems to be the

Bethanney Janney. J (✉) · J. Premkumar
Department of Biomedical Engineering, Sathyabama Institute of Science and Technology, Chennai, India

S. Emalda Roslin
Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, India

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2021

A. K. Manocha et al. (eds.), *Computational Intelligence in Healthcare*, Health Information Science, https://doi.org/10.1007/978-3-030-68723-6_14

AIP Conference Proceedings

Data analytics to find out the effect of firework emissions on quality of air: A case study

Cite as: AIP Conference Proceedings 2201, 020008 (2019); https://doi.org/10.1063/1.5141432
Published Online: 17 December 2019

R. Subhashini, B. Keerthi Samhitha, Suja Cherukullapurath Mana, and Jithina Jose



© 2019 Author(s).

Data Analytics to Find Out the Effect of Firework Emissions on Quality of Air: A Case Study

R. Subhashini[a)], B. Keerthi Samhitha[b)], Suja Cherukullapurath Mana, Jithina Jose

School of Computing, Sathyabama Institute of Science and Technology, Chennai, India.

[a)]subhagopi@gmail.com

[b)]samhitha711@gmail.com

Abstract. This paper studies about the impact of firework emission on air quality of Chennai. This paper considers the Diwali festival season for the studies as it is the time when a lot of burning of firework is happening. Researchers have observed that due to the heavy use of fire crackers the air quality during the Diwali festival season is dangerously low in comparison with other times of the year. . The across the board consuming of fireworks discharges immense measure of harmful gases and colossal measure of pressurized canned products and particulate issue. The present examination canters around the impact of firecrackers on encompassing air quality amid Diwali Festival in Chennai City. In this investigation, PM 2.5, SO2, NOx, CO and O3 were assessed at two agent areas, for pre Diwali, Diwali day and days following the festival. The dimensions of these contaminations saw amid Diwali were observed to be higher when contrasted with Pre Diwali. Consequently it came to a most extreme fixation upon the arrival of the celebration and afterward the qualities steadily diminished after the celebration. The outcomes demonstrate that firecracker outflows amid Diwali celebration influenced the surrounding air quality hurtfully because of emanation and increment of PM2.5, SO2, NOx, CO and O3.
Keywords. Air Pollution, Air quality monitoring, Firework emissions, Data Analytics

INTRODUCTION

Diwali is a celebration of lights which is praised each year amid October/November with extraordinary energy in all over India. Firecrackers release follow gases and particulates into the air, which produce thick billows of smoke, convergence of which relies upon the structure of saltines and sparklers (Barman et al., 2008). A lot of research has been made based on this. Ying Wang et al., (2007) investigate the influence of like particulate matter and its components during various festivals. In one of the paper In another paper Ravindra et al. (2003)[8] detailed that firecrackers cause momentary variety in quality of air and watched multiple times increment in PM10 and TSPM fixations in Hisar city (India) amid Diwali celebration. The investigation completed by Attri et al., (2001) demonstrated the arrangement of ozone because of the consuming of shimmers amid Diwali in Delhi. In Indian condition, ponders have additionally been accounted for broadcasting live quality debasement for the firecracker exercises at the season of Diwali celebration however they are just a few. Despite the fact that they burst the crackers for three days prior, amid and after Diwali, enormous measure of wafers and shimmers are scorched fundamentally at night on the Diwali day. The broad utilization of firecrackers was seen to be lead to momentary distinction in purity of air . The dimensions of these air toxins saw amid Diwali were observed to be genuinely high.

METHODOLOGY

A. Study Area

Chennai, a coastal area on the Bay of Bengal in eastern India. The rise over the ocean level is 6 m. The city region is around 70 sq. miles and the metro (by and large) region is around 456 sq. miles. The Metropolitan Chennai covers with three locale (Chennai, Kanchipuram and Thiruvallur). It encounters a tropical atmosphere. In this manner the climate is ordinarily damp and hot. There is just a little variety between the real season's viz. summer, winter and rainstorm because of the nearness and area to the Indian Ocean. The distinctive reasons for air contamination in Chennai City are fundamentally because of the accompanying classifications, for example, Transport, Industries, Residential and Others. Chennai can be partitioned into four regions, North, Central, South and West. The Northern part is generally a mechanical zone including petrochemical enterprises in the Manali region. The sample site chosen are the residential areas of south Chennai such as IIT (Indian Institute of Technology) and Alandur. Monitoring of

Second International Conference on Materials Science, Smart Structures and Applications
AIP Conf. Proc. 2201, 020008-1–020008-6; https://doi.org/10.1063/1.5141432
Published by AIP Publishing. 978-0-7354-1952-0/$30.00

Low-Cost Bone Mineral Densitometer 191
Riddhi Vinchhi, Neha Zimare, Shivangi Agarwal, and Bharti Joshi

Smart Infusion Pump Control: The Control System Perspective 199
J. V. Alamelu and A. Mythili

Automated Detection of Normal and Cardiac Heart Disease Using Chaos Attributes and Online Sequential Extreme Learning Machine 213
Ram Sewak Singh, Demissie Jobir Gelmecha, Dereje Tekilu Aseffa, Tadesse Hailu Ayane, and Devendra Kumar Sinha

Interference Reduction in ECG Signal Using IIR Digital Filter Based on GA and Its Simulation 235
Ranjit Singh Chauhan

Contactless Measurement of Heart Rate from Live Video and Comparison with Standard Method 257
A. N. Nithyaa, S. Sakthivel, K. Santhosh Kumar, and P. Pradeep Raj

Automatic Melanoma Diagnosis and Classification on Dermoscopic Images 271
Bethanney Janney. J, S. Emalda Roslin, and J. Premkumar

GI Cloud Design: Issues and Perspectives 287
V. Lakshmi Narasimhan, A. K. Sala, and Anne Shergill

A Hybrid Method for Detection of Coronavirus Through X-rays Using Convolutional Neural Networks and Support Vector Machine 305
P. Srinivasa Rao

Feature Extraction Using GLSM for DDSM Mammogram Images 317
Neha S. Shahare and D. M. Yadav

Deep Learning-Based Techniques to Identify COVID-19 Patients Using Medical Image Segmentation 327
Rachna Jain, Shreyansh Singh, Surykant Swami, and Sanjeev kumar

Emerging Trends of Bioinformatics in Health Informatics 343
Mahi Sharma, Shuvhra Mondal, Sudeshna Bhattacharjee, and Neetu Jabalia

Computational Methods for Health Informatics 369
Jayakishan Meher

Computational Model of a Pacinian Corpuscle for Hybrid-Stimuli: Spike-Rate and Threshold Characteristics 379
V. Madhan Kumar, Venkatraman Sadanand, and M. Manivannan

Index 397

International Conference on Communication and Signal Processing, April 4-6, 2019, India

Automatic Fire Rescue System using IoT

T. Sahithi, T. Pranathi and A. Pravin

***Abstract*—The fire rescue system provides complete help to the lonely houses, in case of fire accidents. In India there are so many lonely houses between 5km, and they don't have an option to ask help from others. During that time our project helps to pass an intimation or message to near fire stations. As our project using the help of pic18 microcontroller we can add sensors which will give the intimation to the nearest fire station. The sensors will have a fixed threshold, when the temperature increases above a fixed threshold then alarm will be activated. As the alarm activates the intimation will go to the nearest fire stations with the help of MIT app. When sensor pass the sensed data to thepic18 board after that, the data will be uploaded, and that is for free to the IoT. Wi-Fi module is used as a wireless communication network in our project. If in the case. sensor is not working in that time house owner can make use of a panic or emergency buttons.**

***Index Terms*—Pic18 Microcontroller, MIT App, Wi-Fi Module, Panic Button, Emergency Button.**

I. Introduction

NOW a days, IoT has become the partner in our daily lives. So, by using IoT we can produce and develop many applications like for safety and predictions. The internet of Things is the devices which connect to the network such as vehicles, and other appliances which are used daily in our life bases that consists of electronics, software, connectivity and actuators which allows them to connect, exchange and interact data. IoT is extinguished in various ways like agriculture, medical, business, research purposes and so on. Usage of IoT makes our lives not easy IoT place a very important role in this present modern world, without sensors then it will be very difficult for some applications to be done Fig. 1. IoT places a crucial role in providing safety issues. It is developed to very large extent in security and safety matters. So, in every application IoT is majorly used and it is cost effective according to the purpose it is used. The IoT is connected to many other things like systems, desktops, smart phones, tablets and phones of traditionally dumb or non-internet devices and daily objects. This is very easy way to know data by sitting at our places and this helps people a lot and reduces their work. The rest of this paper Section II discussed about the related work. Section III and IV describes the proposed system and algorithm. Results are discussed in Section V and finally concludes the paper in Section VI.

II. Related Works

NRF2401 is a wireless communication for short range. For long distances GPRS will be used. Wireless multi sensor and fire sensor are used. Fire probability will be judged by the BP algorithm. Cost is very low, and alarm false rate is low. For fire breaks it uses single parameter. ARM9 is used to made centre console. It uses low budget and very easy to build. S3C2$$0A ARM is embedded using BP algorithm. It uses is mainly multi user wireless sensor network [1]. There is a huge loss in property in people lives. Fire alarm equipment placing is necessary, to reduce fire in building automatically. This paper mainly aims on principle and composition. Information will be retransmitted to fire controller [2]. The project is made for warehouses using

An overview on Change Detection and a Case Study Using Multi-temporal Satellite Imagery

N. Anusha [1,*],
Research Scholar,
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
OMR Road,
Chennai, India
anusha.nallapareddy@gmail.com

B.Bharathi[2]
Professor,
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
OMR Road,
Chennai, India
bharathivaradhu@gmail.com

Abstract*—Satellite imagery based change detection plays an important role in analyzing the after effects of natural disasters, detecting the changes in city limits due to rapid urbanization, updating the map database, monitoring the factors impacting agriculture, etc., The remote sensors mounted on satellites or aircrafts absorb the light reflected by the earth's surface. The output of these sensors will be a digital image which represents the scene being perceived. In order to extract the useful information from these images, various image processing techniques need to be employed. In this paper, a detailed outline of the steps and various techniques used for detecting the changes in multi temporal remote sensing images is discussed and a case study is done by taking multi-temporal Landsat-8 images covering Hyderabad city. Image differencing method is applied in order to find the changes in the Hyderabad city limits over 2013December and 2017 December time periods.

***Keywords*— Change detection, Remote sensing, Urbanization, Region of Interest (ROI), Filtering**.

I. INTRODUCTION

Remote sensing refers to gathering or sensing of information about an object without making any physical contact [1]. Remote sensing has many applications in various fields like military, intelligence, commercial, economic, land surveying, humanitarian, geography, etc., Geo Informatics System (GIS) is the most important remote sensing technique. Primary factors in RS are energy source, wavelength and interaction mechanisms. The data collected by the RADAR during floods will be helpful, as it is not affected by atmospheric interference and cloud cover. Optical remote sensing utilizes visible, short-wave infrared sensors to capture the images of different parts of earth. In remote sensing, the Area of Interest (AOI) is captured with help of a satellite or aircraft or from a certain height from the ground level. It takes considerable time to grasp or analyze the damages/changes with human eye. Therefore the solution is digital image processing.

Image Processing includes image enhancement, image degradation, image segmentation, image classification. This has vast application in the field of medical imaging, military, biometrics, satellite imagery, etc. Image refers to interpretation of images[2]. In remote sensing system temporal resolution is the imaging of the exact same area from the same viewing angle at different periods of time. Sensor's temporal resolution depends on factors like swath overlap, latitude and sensor/satellite capabilities. For example, when a natural disaster like earth quake/flood has occurred, then how much area is affected, how much loss occurred and what measures can be taken to help the people in affected area. Detecting changes in the features of the spectral characteristics only can be done by frequently collecting and comparing the imagery at different intervals of time [3].

Using logistic regression addressed techniques for segmentation and supervised classification of multitemporal multispectral SPOT 5 satellite images. Based on the heights of the buildings, the changes were found. This method fails to detect if any new buildings are constructed with same height or sharp edges, mostly suitable for industrial areas [4]. Conditional Random Fields (CRF) model which is a supervised method is used for contextual classification. The dataset/images taken for experimenting are limited to medium resolution [5]. Using unsupervised classification proposed a framework for detecting changes in buildings by utilizing the height and spectral features information. The method applied is a complicated method. This method failed to detect few buildings and changes present in those undetected buildings [6].An automatic change detection method is proposed to detect changes in high resolution multispectral images collected by ZY-3.Two threshold groups are selected instead of using single threshold point in order to enhance the quality of training samples [7]. Hierarchical Change Detection Tests (HCDTs) which is one of the effective online algorithms is presented. It is composed of two hierarchical layers, detection layer and validation layer, which help for detection of changes in data streams. While segmenting the images, few buildings are lost in the process [8].A novel change detection method called super pixel-based change detection method is addressed for extracting features from the very high resolution multi-spectral image[9]. The Table I below gives the Comparison of previous studies.

An Intrusion Detection System using Opposition based Particle Swarm Optimization Algorithm and PNN

T. Sree Kala,
Department of Computer Science
Bharathiar University,
Coimbatore, India
sreekalatm@gmail.com

A. Christy
Department of Computer Science,
Sathyabama University
Chennai, India
ac.christy@gmail.com

***Abstract* - Network security became a viral topic nowadays, Anomaly-based Intrusion Detection Systems [1] (IDSs) plays an indispensable role in identifying the attacks from networks and the detection rate and accuracy are said to be high. The proposed work explore this topic and solve this issue by the IDS model developed using Artificial Neural Network (ANN). This model uses Feed - Forward Neural Net algorithms and Probabilistic Neural Network and Oppositional based on Particle Swarm Optimization Algorithm for lessen the computational overhead and boost the performance level. The whole computing overhead produced in its execution and training are get minimized by the various optimization techniques used in these developed ANN-based IDS system. The experimental study on the developed system tested using the standard NSL-KDD dataset performs well, while compare with other intrusion detection models, built using NN, RB and OPSO algorithms.**

Keywords: ANN, Feed Forward Network, Probabilistic Neural Network, OPSO, NSL KDD dataset.

I. INTRODUCTION

Many organizations are piercing to secure their networks with growing techniques. Even though there are some preventive measures like encryption, authentication, and firewall, still attackers are able to find a way for invisible attacks recorded agent in the system Performance analysis of the benchmark NSL-KDD dataset [4], on the proposed system shows that it achieves good result while matched with other intrusion detection models built using NN, RB and OPSO procedures.

II. REVIEW OF LITERATURE

Ryan et al [6] discussed about an IDS based on anomaly detection developed using back-propagation Multi-Layer Perceptron (MLP) to detect typical user's profile. This MLP model checking for the intrusions by using the users commands in every final session log. The developed model has three layers including two hidden layers and it detects 22 cases out of 24 attacks. But it has a disadvantage that will not identify the new and anonymous attacks.

Cannady et al. [7] newly developed a intrusion detection system that has a three layer Neural Network model to classify the normal or misuse class in the network connection. The nine feature vectors extracted from the data set using the commands and connections in their proposed work. P.L.Nur et al. [8] and K. Labib at al [9] were discussed. Intrusion detection systems built using Self- Organizing Maps (SOM).

An Intelligent Computational Model To Predict Target Genes For Infectious Disease

D. Narmadha
Department of CSE
Sathyabama Inst. Of Science and Technology
Chennai, India
ggnaveengg@gmail.com

Dr. A. Pravin
CSE Department
Sathyabama Inst. Of Science and Technology
Chennai, India
pravin.cse@sathyabama.ac.in

***Abstract*— Biological databases consist of variety of information such as the interactions between proteins, biological functions of proteins and the structure of the proteins. Most of the biological functions of the human body are coordinated by the interactions of proteins with other proteins. Proteins don't act alone. It interacts with other proteins in a biological network to perform numerous functions such as signaling, gene expression. Most of the biological functions are performed with the help of protein interaction network. There is a great need to understand the protein interaction network that helps the researcher to identify proteins associated for various diseases. Many intelligent based computational based approaches have been analyzed by the researchers to improve the prediction exactness of the vital proteins for diseases. However, it is challenging to predict the vital proteins associated with diseases. In this research, we outline the traditional data mining and machine learning methods available to predict the proteins associated with the disease, projecting the strengths, weaknesses and challenges of each model. Furthermore, a new research direction to mine the most vital proteins is proposed and results are remarkable than traditional methods**

Keywords— ***Protein-Protein Interaction, Graph theory, Graph Coloring, Genetic Algorithm***

I. INTRODUCTION

Genes or proteins perform a very central responsibility for the existence of the living organisms. The study about protein-protein interaction network has attracted researchers to identify the proteins that are associated with specific diseases. The task of identifying most vital proteins is not only needed to understand the requirement for the existence of the organism, but also to develop the novel complex centrality measure to find vital proteins. It uses a combination of protein complexes and topological features. If the protein occurs more frequently, there is an increase in the probability of finding essential proteins. The authors in [6] gave insight into the combination of local and global properties to identify indispensable proteins. It combines the topological properties such as local density, betweenness centrality and in-degree centrality measures. It considers all the direct and indirect connections to the target protein. The authors of [7] introduced topology based method for target protein prediction. As the result insight provided by the researchers the influence of the protein decreases with increase in the distance between proteins. The researchers in [8] introduced local interaction density based method to find essential proteins. It evaluates the relationship between proteins and its interacting proteins. The researchers [9] gave research insight to predict the most influential node in the network by calculating the degree of each node and the total degree of the complete protein network taken for consideration. The researchers of [10] introduced CoEWC (Co-Expression weighted by Clustering coefficient) method that combines the structural arrangement of proteins and biological properties for target protein prediction. Pearson Correlation Coefficient is used to explore the degree of correlation between proteins. The methodology described in [11] introduced prior knowledge based method to find essential proteins. It starts with known set of target proteins and for each of the neighboring proteins of the known target proteins; select the one that has high probability of becoming a target protein and update the target protein set.

In recent times, computational approaches such as machine learning model have also gained the attention of

An Enhanced Method for Topic Modeling using Concept-Latent

Christy A.
Faculty of Computing
Sathyabama Institute of Science and Technology
Chennai, India. ac.christy@gmail.com

Meera Gandhi G.
Faculty of Computing
Sathyabama Institute of Science and Technology, Chennai, India
professorgandhi29@gmail.com

***Abstract*- In this digital era, Topic Modeling plays a vital role in clustering of documents. In topic modeling, each document is represented as a collection of words. Labeling appropriate word distributions to the defined topic is a challenging task. Topic modeling using unsupervised algorithm does not align topics with those as it exists in the real world metaphor. This problem of aligning topic with real world objects can be carried out by two different methods – either to use a supervised learning algorithm that assigns individual topic with a label acquired from existing data or to use a predefined set of topics and to match them with the distribution of words. The above said methods have some drawbacks as the former method may match a label which does not convey a semantic meaning with the word distributions whereas the latter lacks the knowledge required to detect the topic from the closed set of topics. In this paper an algorithm called Concept-Latent Dirichlet Allocation (Concept-LDA) is proposed which incorporates reinforcement learning for topic modeling which inturn improves the accuracy of resulting topics and of the topic labeling. Experiments were conducted using LDA and Concept-LDA for topic modeling and it is shown that Concept-LDA showing better accuracy than LDA.**

Keywords - Documents, Reinforcement learning, Latent Dirichlet Allocation, Gamma distribution, Topic modeling, etc.

I. INTRODUCTION

In this era of Information Technology, information exists

II. LITERATURE SURVEY

Unsupervised feature selection algorithm is very challenging since features are selected by adopting the method of clustering which satisfies a particular criterion.

Anna Drummond et al (2013), have built a topic model and the topics are distributed as a set of features describing the documents. The resulting vectors thus obtained can be then used with algorithms to perform clustering [1]. In Vanilla LDA for unsupervised clustering ɵ vectors are taken input. LDA uses the concepts of Dirichlet on the input ɵ vectors and organize the mixture of words and topics present in the collection of document.

Ben D. Fulcher and Nick S. Jones (2014) have solved time series classification using feature selection approach; in which the features are identified in scientific time series analysis literature. Datasets were derived from UCR Time Series Clustering or Classification Homepage and a greedy forward feature selection was adopted [2]. Derek Greene and P´adraig Cunningham (2006) have narrated that the performance of SVM reduces in the case of supervised learning algorithms, where as the performance increases along the diagonal entries.

Chaitanya Chemudugunta et al (2008) proposed a concept topic model in which a multi nomial distribution with

Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019)
IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8

AN EFFICIENT OBJECT DETECTION MODEL USING CONVOLUTION NEURAL NETWORKS

Dr. Ulagamuthalvi [1] ,J.B. Janet Felicita[2] ,Abinaya D[3]
[1] Associate Professsor, Dept of CSE, Sathyabama Institute of Science & Technology, India
[2] Student, Dept of CSE, Sathyabama Institute of Science and Technology, Chennai, India
[3] Student, Dept of CSE, Sathyabama Institute of Science and Technology, Chennai, India

Abstract **— Image processing and computer vision have gained an enormous advance in the field of machine learning techniques. Some of the major research areas within machine learning are Object detection and Scene Recognition. Though there are numerous existing works related to the specified fields object detection still encounters numerous challenges when it comes to implementing in the real-time scenario. The problem occurs in the detection due to various objects present in the background. Object detection mechanism detects a specified object when a particular scene is given. Classifiers like SVM and Neural Networks are used to train the classifier in such a way they are able to detect an object when a new image is given. In this paper, we have proposed a model which detects texts from an image. Bounding boxes are used to detect the texts and localize it. The neural network is used to train the model where numerous images having texts are given as the training set. The performance evaluation is done on the model and it is observed that it detects the texts when a new image is given. Object detection is a fundamental problem in computer vision, which aims to detect general objects in images.**

Keywords— ***Object Detection, Bounding Boxes, SVM, Neural Networks, Classification, Texts, Prediction***

I. INTRODUCTION

Numerous research works are taking place across the globe in various fields. One such field is Machine Learning techniques that are catching up a wide variety of research works especially in object detection and scene recognition [1]. Object detection plays a vital role in detecting and identifying objects in a scene or an image such as it is used in identifying the number of vehicles in a particular time on a road [2] or the number of passengers standing in a bus stop during a peak time [3] and a lot more. Whereas scene recognition plays a very immersive role when it comes to classifying flood-affected areas [4] and the land cover classification [5] in a particular region. The detection is done using various classifiers such as SVM classifier [6], KNN classifier [7] and now a day's Neural Networks are playing a huge role in detecting the objects [8].

Object detection is done on an image where the objects are first trained on a various group of objects and then the classifier comes to a stage where when a new image is given it can automatically detect all the trained objects which more accuracy by bounding boxes around it and naming the label with which it had been trained. Previous studies have shown that the use of SVM classifier in detecting objects and bounding it with text boxes [9]. Numerous researchers have even use Convolution Neural Networks to extract the features of the trained objects and use these feature to classify the new objects [10].

ALERT SYSTEM FOR DRIVER'S DROWSINESS USING IMAGE PROCESSING

Ratna Kaavya M[1]
Student,
Department of Computer Science
Sathyabama Institute of Science and Technology,
Chennai-600119,
kaavyamuthukumaran@gmail.com

Ramya V[2]
Student,
Department of Computer Science
Sathyabama Institute of Science and Technology,
Chennai-600119,
ramyav6698@gmail.com

Ramya G Franklin[3]
Assistant Professor,
Department of Computer Science
Sathyabama Institute of Science and Technology,
Chennai-600119,
mikella.prabu@gmail.com

Abstract **- Accidents occur all over the world cause of being not able to concentrate on the road while driving. The concentration is missed due to driving the car without resting which makes the person drowsy. And this drowsiness becomes the reason for major accidents. This problem is overcome by developing the various system in detecting the sleep. The proposed system here uses the Raspberry Pi and various sensors like Gas Sensor, Vibration Sensor for the detection of the type of drowsiness. The driver is been monitored by placing a camera which captures the vital sign. If the eye is closed for a longer period of time then the image of the person is sent. The accident is detected using a vibration sensor and the server is notified by sending latitude and longitude. The location of the car is sent by the IOT modem which is embedded in the car. If the driver consumed alcohol it is sensed by the gas sensor and the server is notified by a message. The motor of the car is continuously running can be stopped or cut-off if the server side is notified about the driver that the person is not in a situation to drive the car. By this, the accident rate is reduced and the risk of customer life is decreased.**

Keywords: Image Processing, Drowsiness Detection System, Fatigue, Raspberry Pi

I. INTRODUCTION

Every human being needs sleep but a lot of people underestimate sleep and it is a recovery time for humans to gain energy and get back to work on a daily basis. Lack of sleep leads the body to work ineffectively and mind to a drowsy state. Lots of report says that 20% accidents due to fatigue-related. Driver fatigue is caused by four factors namely sleep, work, time and physical. Every four minutes one death occur due to accident. To detect drowsiness many technique has been used eye retina detection, pulse pattern detection. Human sleep can be parted as into three things: Fully Awake, Non Rapid Eye Movement and Rapid Eye Movement sleep and the Anti-fatigue system into two things: One is to capture the image of the driver and estimate the drowsiness, Second is to estimate drowsiness based on heart rates [1]. According to study the drowsy state occurs due

978-1-5386-9353-7/19/$31.00 @ 2019 IEEE

Adaptive Cross Layer Scheme for Multimedia Transmission over Wireless Sensor Networks

Jenila L, Research Scholar
Computer Science And Engineering
Sathyabama Institute Of Science And Technology
Chennai, India
jenilacse@gmail.com

Aroul Canessane R, Professor
Computer Science And Engineering
Sathyabama Institute Of Science And Technology
Chennai, India
aroulcanessane@gmail.com

***Abstract*—WSN have a wide choice of applications also plays a major responsibility in raising a timely and precise information. The information that is sensed by the sensor nodes should be delivered to the destination in a timely manner without any unnecessary delays and losses. Cross layer Scheme for Multimedia Transmission (CLSMT) for wireless sensor network is proposed to improve the network efficiency. Data overloading and data latency can be reduced by using cross layer mechanism. Multimedia transmissions include larger set of data transmission that obviously consumes huge amount of energy. Therefore the nodes can be defined with predefined threshold values for energy level, packet scheduling and bandwidth allocation. Therefore this improves the network efficiency in terms of delivery rates also increases network throughput.**

Keywords— Cross layer scheme, multimedia transmission, Parametric Thresholds, Wireless Sensor Networks.

I. INTRODUCTION

In several applications WSN plays a major role in monitoring and controlling the wireless environments. Video application comes under multimedia is one of the major applications in sensor networks that includes video conferencing, video chatting, video surveillance in day-to-day life. Other than scalar data, for example, weight, humidity and temperature, sensor hubs might be invested with multimedia detecting units, improving observing applications with sound, pictures and video streams as well.

Cross layer design handles transmission for multimedia over wireless networks, it solves global optimisation of system performance [1]. Direct communication between non-adjacent layers, allowing sharing information across layers, merging adjacent layers or creating new interfaces between layers are well developed by applying cross layer mechanism.

Multimedia information which is transmitted in WSNs is to be sure a difficult task, either because of the huge measure of information to be transmitted or because of the continuous idea of some mixed media observing applications. The system lifetime straightforwardly relies upon the energy utilization rate, and interactive media gushing is required to expend considerably more energy than transmission of normal information [2], [3].Packets or frames conveying bits of any signals like sound, picture, and video information from various source hubs might cross the system toward the sink, regardless of whether sources' transmissions are nonstop or activated based [4]. As packets transferring uses more energy and transmission mutes might be debilitated because of energy exhaustion in half of the transmission process, therefore the aggregate sum of data transmitted over the system ought to be limited however much as could reasonably be expected, conceivably dragging out the system lifetime.

II. RELATED WORKS

Numerous research works have been carried out for video transmission over wireless networks and to produce better throughput at the receiver end. The design for video transmission from a sensor to the UMTS terminal was proposed requiring little to no effort and power for the reason that the video sensors can be inside the range of UMTS for best usage of UMTS inclusion [5]. The joining of the portable system framework and sensor impromptu systems can decrease the expense of structure new foundation and empower the huge scale arrangement of sensor systems. Braided Cooperative Reliable transport (BCRT) calculation for dependable video transmission inside mine passage was proposed [6]. This BCRT scheme keeps up one-bounce unwavering quality for guaranteeing start to finish dependability so it is strong to way breaks. Also, the BCRT-based WSN utilizes three continuous forward neighbors to change the information rate for improving the detecting dependability. For productive multi-path directing of ongoing video streams, a path distribution calculation [7] was proposed to fulfill the postpone requirement of video streams while adjusting energy and data transfer capacity utilization among all the accessible routes. For the situation that the total data transfer capacity is as yet insufficient to fulfill the required coding rate, as well as cross-layer method for versatile coding as per way status is included.

An application-centric routing framework [8] was proposed for real-time video transmission in multi-hop wireless networks, where expected video distortion is adopted as the routing metric. Also a quality-driven cross-layer optimization is applied for enhancing the flexibility and robustness of path direction using joint optimization of path selection and video coding, subsequently augmenting the client saw video quality under a given video playback defer limitation. Here both Routing and rate control combined in a unified convex optimization [9] detailing was proposed alongside circulated joint arrangement dependent on cross-layer plan. Firstly twisting model was built up that catches both the effect of encoder quantization and casing misfortune event of system blockage on the general video quality. At that point, the ideal joint rate control and steering plan is acknowledged by adjusting its rate to the time-shifting traffic and limiting the blockage rate. Novel scheme named bandwidth sensitive traffic [10] was proposed for entrepreneurially sending information frames and give transfer speed assurance to the transmission capacity touchy traffic. The normal expected bandwidth capacity (EAB) and the expected transmission cost (ETC) under OR operation are evaluated dependent on the neighbourhood accessible

978-1-7281-5286-8/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: Middlesex University. Downloaded on July 20,2020 at 07:11:21 UTC from IEEE Xplore. Restrictions apply.

Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019)
IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8

A Sustainable Vehicle Parking using IoT

Anandhi T
Assistant Professor
Department of Computer Science and Engineering
Sathyabama Institute of Science and Technology
anandhitamil2908@gmail.com

Kishore Kumaar V.S., Maha Ganesh S.
U.G student
Department of Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
kishorevsrini@gmail.com,,mahaganesh97@gmail.com

Gomathi R.M
Assistant Professor
Department of Information and Technology
Sathyabama Institute of Science and Technology

Abstract—**we present a highly efficient method for vehicle parking management in large scale parking lots based on IoT .The proposed system can be installed in a Raspberry pi module which will help us in solving the current problems in vehicle parking management. The system consists of an on-site deployment of an IoT module that is used to monitor and signalize the state of availability of each single parking slot. This can be done on a mobile phone using android application. This application can be managed by an administrator at the malls. This project primarily makes finding of parking systems more efficient and less time-consuming, and also eliminates the need of staff to be employed for such purpose of these places, thus increasing profit margins. The system helps a user know the availability of parking slots on a real time basis. Android mobile application is used to display the total parking slots in the parking lot, number of parking slots available, occupied parking slots and the reserved parking slots. To book the parking slot at that time your information are encoded to the QR code. This makes our parking system more secured. Once you park the vehicle in the slot, the timer will start to monitor the parking time. If time exceeds alert message is sent to the user's mobile.**

Keywords—Mobile Application, Internet of Things, Raspberry, SMS

I. INTRODUCTION

In the current world, the issue of finding a parking slot without a huge amount of waiting time has become the most difficult thing when it comes to parking. It has been noted that a huge amount of a person's time is being wasted in finding a perfect parking slot for his vehicle. On a public holiday, things become worse all the major shopping malls and commercial areas are full with the parking slots and the people have to wait for more time to get a parking slot. Parking Manager is an idea which when implemented can solve these kinds of issues. This technology helps in checking the availability of parking slots and helps further by providing an interface to book parking slots in advance before the actual time of entry like the similar manner in which we do for hotels. The primary goal of this project is to make people do parking in a hassle free manner with the help of latest technology.

II. RELATED WORK

[1] Nikhil Palde, Chhaya Nawale and Sunita Kute
As production of vehicle grows and the global population grows, more space is needed today. The project provides a new parking system, parking system and the Android approach, so that the driver can find free parking for short periods of time. We review and compare different technologies to determine the best technologies for developing this system. The system uses an infrared sensor to detect the presence of a vehicle in a parking lot and display empty space. We create a separate application for the smartphone and use this application to find an easier and easier way to get to your destination using the Bluetooth module. The Android approach to the Parkhaus system includes automatic parking, free parking signs and free parking for smartphone apps. This project describes how to properly and effectively use the parking system from the entrance to the entrance.

[2] Karma Tsheten Dorjee, Deepak Rasaily and Bishal Cintury
This project is a revolutionary electronic money system that provides the best solution for car drivers, city and private

A Study on Various Semantic Metadata Standards to Improve Data Usability

Suja Cherukullapurath Mana
Sathyabama Institute of Science and Technology
Chennai ,India
cmsuja@gmail.com

T.Sasipraba
Sathyabama Institute of Science and Technology
Chennai,India
prove@sathyabama.ac.in

***Abstract*— An efficient meta data systems will helps to improve the usability of data. Metadata will describe the data being stored and it will helps in adding more value and usability to the stored data. Data retrieval also will be made easy by using an efficient metadata system. This survey paper will study some of the metadata standards and perform a comparison based on the usability of each metadata standards.**

Keywords—data usability, data retrieval, semantic web, metadata

I. INTRODUCTION

As the volume of data getting generated day by day is becoming huge, efficient methodologies to store and manipulate data is becoming a prominent area of research. Big data and data warehouse systems are in raise to handle this huge volume of data. As these systems will have heterogeneous data, efficient retrieval of data is also a point of concern. Meta data system can help in solving some of the heterogeneous data storage and retrieval problems. Metadata describes the data being stored in data storage. Efficient metadata systems can contribute in solving some of the data storage and retrieval problems. An efficient metadata system can also contribute in improving data quality, supporting data integration and maintenance process. It will also help in clarifying data lineage. This paper will perform a survey on some of the ontology based meta data standards and also perform research on how these meta data standards can be efficiently utilized to create a meta data system to increase the usability of data.

II. AN OVERVIEW OF RESOURCE DESCRIPTION FRAMEWORK(RDF)

Resource description frame work (rdf) is the key framework applicable for semantic web [2]. It is an efficient framework used to describe various recourses on the web. RDF will define data in a triple format namely subject, object and predicate. A RDF graph includes set of these triples. A subject is what being described and predicate indicates what property of the subject is being described [4]. Object is what is being described [4]. An important concept of RDF is the classes.

Rdf resources are grouped into classes [13]. Members of the classes are called resources. One class can be a subclass of another class [6]. This scenario can be described by the property rdf: subclassof .Some of the class descriptions are given below [13].

Resourse – All things described by rdf are qualified to be a member of this class.

Class- This is a collection of resources which are classes.

Literal– This class include all literals . Strings , integers are some examples of literals. Litera is a subclass of resources.

XMLliterals- All literals used in XML are part of this class. This class is a subclass of literals.

The relation between subject resources and object resources are described by properties [13]. The property rdf: subpropertyof is used to define the concept that one property is a sub property of another [17]. Similarly rdf:range is used to describe the fact that one property can be instance of one more classes. Rdf:label is also an instance of rdf:property and it is used to describe label of a resource[13].
RDF/XML is the language used in Rdf. Rdf recourses are identified by web uniform identifiers. A sample rdf description is given below.

```
<rdf:Description
rdf:about="http://www.bookseller.com/c/Databasesystems"
>
  <c:author>Shilbershatz</c:author>
  <c:country>USA</c:country>
  <c:edition>Sixth</c:edition>
  <c:price>100</c:price>
  <c:year>1990</c:year>
</rdf:Description>
```

RDF representation can be represented as a RDF graph. A sample RDF graph describing a person is given below.



Fig 1: A simple RDF graph

978-1-5386-9471-8/19/$31.00 ©2019 IEEE

A Review on Internet of Things Protocol and Service Oriented Middleware

Y. Justin Dhas and P. Jeyanthi

***Abstract*—This paper surveys a review of Internet of Things (IoT) protocols, Service oriented Middleware in IoT. The modern development of IoT, expected to create many divorce application in health care without human intervention. Various protocols are involved in the applications development. Researchers are doing research for desirable protocol with all functionalities. Middleware for an IoT provides interoperability between the devices or applications. The engineering of an IoT dependent on Service Oriented Architecture (SOA), it operates as middleware. We survey the existing SOA based IoT middleware and its functionalities.**

***Index Terms*—Internet of things, IoT Middleware, IoT Protocols and IoT Security**

I. Introduction

THE evaluation of web 3.0 the Internet of Things (IoT) grows rapidly in the field of service computing. IoT is aims to process the data, and capable to sensing, integrating and analyse the data in all aspects of the world [1].Internet infrastructure is interconnected with many smart objects and devices by using Machine-to-Machine (M2M) technologies [2].M2M refers to technologies that allow both wired and wireless devices can communicate each other.

In a ubiquitous computing like IoT generates large amount of data along with the heterogeneous infrastructure and its applications. For making ubiquitous applications is a challenge task in IoT [3]. This Middleware provides concatenation and interoperability of applications and services functioning on dissimilar computing and communications accessories. These services it provides identifications of objects, authentication, authorization, and security.

The IoT middleware designed for many applications and it's covered various domain, different languages and different infrastructure. SOA middleware provides effective system interoperability. IoT have four different components or pillars like WSN (Wireless Sensor Networks), M2M (Machine-to-Machine), RFID (Radio Frequency Identification), and SCADA (Supervisory Control And Data Acquisition) [4],[5].All the four components of IoT are needed for the full functioning of IoT middleware. In this literature we examine the various protocols and SOA based IoT middleware and further a detailed comparison of security and privacy issues of SOA based IoT middleware is analyzed.

The rest of the paper is organized as follows. The IOT application protocols and SOA middleware are described in Section II and III respectively. Section IV lists the security issues. At last, Section V concludes the paper with conclusion.

II. IoT Application Protocols

The Fig. 1 shows the various IoT application protocols and its transport Layer Protocols

1. CoAP - Constrained Application Protocol
2. AMQP - Advanced Message Queuing Protocol
3. MQTT - Message Queue Telemetry Transport
4. MQTT-SN
5. AMQP - Advanced Message Queuing Protocol
6. DDS - Data Distribution Service
7. HTTP REST

A. Constrained Application Protocol

CoAP is an application layer protocol made by IETF Constrained RESTful Environments (CoRE) [6],[7] for IoT Applications. CoAP is REpresentational State Transfer (REST) protocol over HTTP for communicating data between client and server. CoAP is bound to User Datagram Protocol (UDP).It is most suitable for IoT applications. CoAP aims to provide low power consumption and compatible for tiny devices.

Application Protocol	
DDS	TCP UDP
CoAP	UDP
AMQP	TCP
MQTT	TCP
MQTT-SN	TCP
XMPP	TCP
HTTP REST	TCP

Fig. 1. IoT Application Protocols

Y.Justin Dhas is with the Research Scholar, Sathyabama Institute of Science and Technology, Chennai, India, e-mail: justindhasy@ gmail.com.

P. Jeyanthi is with the Faculty of Computing, Sathyabama Institute of Science and Technology, Chennai, India, e-mail: jeyanthiprabha@yahoo.com.

978-1-5386-7595-3/19/$31.00 ©2019 IEEE

IEEE Advancing Technology for Humanity

A Novel Design and Analysis of Reversible Barallel Shifter Using New FF-Gate

Ranjith. S
Research Scholar,
Sathyabama University, Chennai,
Tamil Nadu, India
Email: ranjithsubramanian90@gmail.com

Dr.T.Ravi
Assistant Professor, Department of ECE,
Sathyabama University, Chennai,
Tamil Nadu, India
Email:ravi_vlsi123@yahoo.co.in

***Abstract*—The need for low power has gained attention towards lot of low power techniques from quantam physics the concept of reversible logic has evolved to reduce the power dissipiation.This paper deals with the design of 8 bit shift register using reversible logic.The existing system the implementation carring out usingFredkin gate Feynman gate.Our design using implementation 4 X 4 ff gate which reduces the gate count as well as the gorbadgeoutput.The proposed design shows better output performance and provides and low power consumption than the existing design**

***Keywords*—barrel Shifter, FFReversible Gate, Fredkin and Feynman reversibleGate.**

I. INTRODUCTION

VLSI(Very large scale intergration) which bring a cluster of transistor in a single chip. The VLSI has been a evergreen technology because of its cmpactablity and flexibility in accompining N no of transistor in a single chip. A standing law which governs the VLSI chip design industry stated by Gordon Moore. The law states that the IC design should be able to intergrate double the number of transistors in the same given area with reducedpower dissipation within the period of 18 months. In fact the doubling period has decreased to a mere 12 months.The number of transistors/gates that can fit into the semiconductor that brieflyexplains the complexity of the functionality that the device can perform. The focus in the area of VLSI research gained a great attention because of its everlasting three golden rules low power, reduced area, and faster devices. This made everything very compactable and digitalized [16].

II. REVERSIBLE LOGIC

In conventional logic the power dissiotation takes place in the form of thermal energy.This can be reduced in the reversible logic.Fromquantam physics it is evident that if a reaction is balanced on both a side than ther is no power loss or any kind of loss in the system.This has been taken as a base of the reversible logic which has N input and N output and in otherword no of inputs and no of output are equal. Reversible logic design should follow the following the condition.1.Fan out and Fan in should be reduced.any one of the output shoulf follow the any one of the input.2.There should be one to one mapping from the input to the output.3.No of gorbadge output should be low as possible[12].A reversible circuit couples reversible gates that do not include loops and fan outs. So the circuits has the equal numbers of input and output wires, each going through an entire circuit.Less numbers of reversible logic gates are used to design a reversible logic circuit and with a less number of constant inputs [6-9].

III. BARREL SHIFTER:

A barrel shifter is often designed as a cascade of parallel 2×1 multiplexers. For a four-bit barrel shifter, two bits are used to intermediate by an intermediate signal, or passes the same data, depending on the value of S[5]. This signal is then shifted by using another multiplexer, which is controlled by S[0].sA (8,3) reversible bidirectional arithmetic and logical shifter are processed all other operations.The barrel shifter performs various operations such as logical right shift, logical left shift, etc. depending on the values of Sra, slow and left control signals.

In the proposed design, the input data is represented as i7, i6, i5, i4, i3, i2, i1, i0 while the

978-1-5090-4855-7/17/$31.00 ©2017 IEEE

A Novel Design and Analysis of Reversible Barallel Shifter Using New FF-Gate

Ranjith. S
Research Scholar,
Sathyabama University, Chennai,
Tamil Nadu, India
Email: ranjithsubramanian90@gmail.com

Dr.T.Ravi
Assistant Professor, Department of ECE,
Sathyabama University, Chennai,
Tamil Nadu, India
Email:ravi_vlsi123@yahoo.co.in

***Abstract*—The need for low power has gained attention towards lot of low power techniques from quantam physics the concept of reversible logic has evolved to reduce the power dissipiation.This paper deals with the design of 8 bit shift register using reversible logic.The existing system the implementation carring out usingFredkin gate Feynman gate.Our design using implementation 4 X 4 ff gate which reduces the gate count as well as the gorbadgeoutput.The proposed design shows better output performance and provides and low power consumption than the existing design**

Keywords—barrel Shifter, FFReversible Gate, Fredkin and Feynman reversibleGate.

I. INTRODUCTION

VLSI(Very large scale intergration) which bring a cluster of transistor in a single chip. The VLSI has been a evergreen technology because of its cmpactablity and flexibility in accompining N no of transistor in a single chip. A standing law which governs the VLSI chip design industry stated by Gordon Moore. The law states that the IC design should be able to intergrate double the number of transistors in the same given area with reducedpower dissipation within the period of 18 months. In fact the doubling period has decreased to a mere 12 months.The number of transistors/gates that can fit into the semiconductor that brieflyexplains the complexity of the functionality that the device can perform. The focus in the area of VLSI research gained a great attention because of its everlasting three golden rules low power, reduced area, and faster devices. This made everything very compactable and digitalized [16].

II. REVERSIBLE LOGIC

In conventional logic the power dissioiation takes place in the form of thermal energy.This can be reduced in the reversible logic.Fromquantam physics it is evident that if a reaction is balanced on both a side than ther is no power loss or any kind of loss in the system.This has been taken as a base of the reversible logic which has N input and N output and in otherword no of inputs and no of output are equal. Reversible logic design should follow the following the condition.1.Fan out and Fan in should be reduced.any one of the output shoulf follow the any one of the input.2.There should be one to one mapping from the input to the output.3.No of gorbadge output should be low as possible[12].A reversible circuit couples reversible gates that do not include loops and fan outs. So the circuits has the equal numbers of input and output wires, each going through an entire circuit.Less numbers of reversible logic gates are used to design a reversible logic circuit and with a less number of constant inputs [6-9].

III. BARREL SHIFTER:

A barrel shifter is often designed as a cascade of parallel 2×1 multiplexers. For a four-bit barrel shifter, two bits are used to intermediate by an intermediate signal, or passes the same data, depending on the value of S[5]. This signal is then shifted by using another multiplexer, which is controlled by S[0].sA (8,3) reversible bidirectional arithmetic and logical shifter are processed all other operations.The barrel shifter performs various operations such as logical right shift, logical left shift, etc. depending on the values of Sra, slow and left control signals.

In the proposed design, the input data is represented as i7, i6, i5, i4, i3, i2, i1, i0 while the

978-1-5090-4855-7/17/$31.00 ©2017 IEEE

International Conference on Inventive Systems and Control (ICISC 2019)
IEEE Xplore Part Number: CFP19806-ART; ISBN: 978-1-5386-3950-4

A DISTINCITVE MODEL TO CLASSIFY TUMOR USING RANDOM FOREST CLASSIFIER

Divya. S[1], R. Vignesh[2], R. Revathy[3]
[1][2][3]Assistant Professor, Department of Computer Science and Engineering
Sathyabama Institute of Science and Technology

***Abstract*—The distinctive machine learning model that was created as a need for doctors/ oncologists who treat patients in critical stages. The present day has many people suffering from cancer where they get diagnosed only during the last stage (4th stage) of cancer. This leads to many untimely deaths of their loved ones for many people. To reduce such risks and provide more effort in saving those lives, this model may be used. This model is made from Random Forest classifier[1] where it classifies a tumor to be either Benign(Non-cancerous) or Malignant(Cancerous). It uses 10 features of tumor subdivided into mean, standard error and worst case value of each to increase its accuracy. The inputs given to this model are obtained from medical imaging and hence do not need any medical tests where time may be wasted. The future of this model relies on the demand where it may lie in being developed into an application or it may be developed into a full-fledged health-care system. The main objective of this model, is to ensure that more time can be bought to save or extend the lifetime of the patient by providing chemotherapy as a preventive measure for an untimely death that may occur. This model predicts with 94.34% accuracy, 93% best case confidence and 56% worst case confidence whether the given data resembles a malignant or benign tumor.**

***Keywords*—** cancer, random forest classifier, oncologist, tumor.

I. INTRODUCTION

The "Tumor Classification model" has been created to overcome the critical situations faced by the medical industry when diagnosing cancer. The model is made using a software called "RapidMiner" which is a professional Data Science software platform [2]. It consists of modifiable operators or functions ranging from data mining to deep learning. The model created using this software has been modified by me to ensure that it reaches a combination of maximum accuracy and minimum error rate. A Model Simulator is included in which the various values are entered to obtain a nearly-accurate output which can be used as a lead. The model can later be integrated into any common programming language like Python or Java to create an app as this software is built on Python and Java. It consists of 10 main features where each is divided into 3 sub-features. The details regarding each will be discussed later in this document. This model is targeted towards not the general audience, but oncologists who suffer to diagnose any 4th stage cancer.

II. DESIGN AND IMPLEMENTATION

The Machine Learning model uses a Random Forest classifier in a software called "RapidMiner". Rapid Miner is the data science software platform which provides an integrated environment for the machine learning, data mining, deep learning and predictive analytics. It is used for various applications as well as for research, educational purpose/training, rapid prototype, and application development and supports all steps of the machine learning process including data preparation, results visualization, model validation and optimization. Rapid Miner is developed on an open core model. The UI of the software is given in Figure 1.



Figure 1: Rapid Miner UI

A. RANDOM FOREST CLASSIFIER

Rapid Miner provides a feature called Auto Model where, when the input data set is given and the relevant options are chosen such that it creates comparisons of all selected approaches in its default modification. When chosen accordingly using the relevant data set [3][4], the top three contenders considering time taken and accuracy where Logistic Regression, Random Forest and Generalized Linear Model. While applying them with modifications in the general UI where I built the software model, the confidence also had to be considered apart from the accuracy and time taken. While considering these three characteristics, Random Forest satisfied them all with an average confidence of 93.33%, 94.34% accuracy and an execution time of 1 second. The

CHAPTER

16

Corrosion behavior of ceramic nanocomposite coatings at nanoscale

A.M. Kamalan Kirubaharan, P. Kuppusami

Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

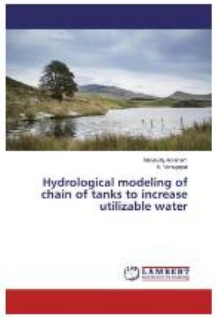
1. Introduction

High-temperature corrosion on materials is one of the major issues for the selection and service life prediction of metallic components, which are operated at high temperatures. For instance, corrosion of mild steel is one of the major issues faced today and leads to huge amount of loss every year. It is necessary to protect the corrosion resistance of metals and alloys because the mechanical and tribological properties are deteriorated at high working temperatures. Inward diffusion of oxygen, sulfur, or carbon could result in precipitation of by-products such as oxides, nitrides, sulfides, carbides, and/or their mixtures that leads to the loss of material stability, which results in the shorter lifetime of the components. Corrosion on alloys can be prevented by (1) the proper choice of material, (2) changing the environment conditions, (3) surface modifications, and (4) cathodic corrosion protection.

Each method of corrosion protection has its own advantages and disadvantages, and the selection of appropriate method depends on the operational conditions of the alloys. For protection against hot corrosion and oxidation, surface modification by the application of overlay coatings is a primary option [1]. Coatings are thin layers of a material deposited onto a surface of the material, as a protective layer to improve the surface properties and to create a protective barrier against corrosion. For technological applications, the coating and substrate requirements may be very stringent. For example, in high-temperature, high-stress, and corrosive environments, the substrate material must have good strength, fatigue, and creep properties. Such substrate materials generally have poor corrosion resistance. In this case, the applied coating must provide protection against corrosion without degrading the metallurgical properties of the substrate. The choice of coating materials may be desirable to protect the underlying substrate materials either as diffusion barrier layers, sacrificial layers, or electrochemical modifiers.

Corrosion Protection at the Nanoscale
https://doi.org/10.1016/B978-0-12-819359-4.00016-7

Copyright © 2020 Elsevier Inc. All rights reserved.



Hydrological modeling of chain of tanks to increase utilizable water

LAP Lambert Academic Publishing (2019-06-19)

€ 61,90

Buy at the MoreBooks! Shop

Scientific approach in assessing the status of existing water bodies and surface water budgeting to identify opportunities for harnessing surface water is carried out along with ways for increasing utilizable water from both surface and groundwater sources. The large regional variation in rainfall and the seasonal character of river run-off in India have prompted human interference with the hydrological cycle. The land use changes taking place in the peri-urban area modify the drainage pattern thereby changing the catchment area of the tanks leading to water stagnation and flooding. Rainfall and runoff events were identified for hydrological modeling so as to assess the validity of USDA-NRCS model. Utilizing the historic data of rainfall for 14 water years runoff was computed for four tanks. Simulation of these tanks have been carried out for a period of 14 water years with the inflows obtained from USDA-NRCS model, estimated demand for irrigation, assessed net losses and connectivity of the tanks in terms of realizing the spill from upstream tanks and return flow from upstream tank ayacuts. The feasibility of increasing utilizable water is discussed in respect of each of the tank.

Book Details:

ISBN-13:	978-620-0-21869-8
ISBN-10:	6200218692
EAN:	9786200218698
Book language:	English

Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Fabrication of p-type cubic γ-CuI by solid iodination process for energy conversion and storage applications

Karthik kumar Chinnakutti [a], Vengatesh Panneerselvam [a], Durai Govindarajan [a], Shyju Thankaraj Salammal [a,b,*]

[a] Centre of Excellence for Energy Research, Sathyabama Institute of Science and Technology, Chennai 600119, India
[b] Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai 600119, India

ARTICLE INFO

Article history:
Received 12 March 2019
Received in revised form 14 April 2019
Accepted 10 May 2019
Available online xxxx

Keywords:
Solid iodination
p-type
γ-CuI
Specific capacitance
Hall measurements

ABSTRACT

A simple solid iodination method is used to achieve an efficient p-type cubic γ-CuI thin film from Cu_3N thin film. The crystallinity and crystal structure of γ-CuI thin films were confirmed by X-ray diffraction (XRD) technique. The maximum specific capacitance of CuI films was found to be 54 mF cm^{-2} at scan rate of 10 mV s^{-1} with excellent cycling stability. The optoelectronic behaviour was studied using UV–Vis–NIR spectroscopy and Hall measurements. These findings suggest that CuI film as a new inexpensive p-type material for third generation solar cells and good electrode for electrochemical energy storage devices.
© 2019 Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the Advanced Materials for Clean Energy and Health Applications (AMCEHA).

1. Introduction

Copper iodide (CuI) is a p-type wide bandgap semiconductor with three (α, β, γ) crystalline phases, preferable diamagnetism, and fast-ion conductivity [1–4]. p-type semi conduction in γ-CuI is due to the presence of excess iodine where it generates holes in the valance band and hence acts as an electron acceptor. Thus, the conducting CuI films (238 S cm^{-1}) with optical bandgap in the range of 2.3–3.1 eV [5], has shown enormous interest and hence being used as a buffer layer in solid-state dye-sensitized solar cells (DSSC) [6–10], organic solar cells [11], and as hole transporting layer in perovskite solar cells [12,13]. Although the CuI acts as a hole collector in DSSC, the device performance showed poor stability which can further be improved by incorporation of thiocyanate ionic liquid into the CuI coating [14]. In perovskite solar cells, CuI has been used as an efficient hole transporting layer. Zhou et al., used CuI layer in ZnPc:C_{60} based bilayer organic solar cells [15]. Kamat et al. employed CuI as an inorganic HTM in a perovskite solar cell and achieved a maximum PCE of 6.0% for a cell-design consisting of a dense TiO_2/meso-structured perovskite/CuI/ gold [12].Sepalage, et al. also reported that, CuI as an inorganic HTM which was deposited by rapid solution-based process, resulted in a maximum PCE of 7.5% [16].

As stated above, CuI have found its potential applications in optoelectronic devices. However, the study of its electrochemical capacitive properties of CuI is scarce. Ezealigo et al. deposited CuI thin film by SILAR method with a maximum specific capacitance of 93 F g^{-1} at scan rate 2 mV s^{-1} [17]. Hence, we directed our research towards the analysis of optoelectronic as well as electrochemical behaviour in order to widen the applications of CuI. Various physical and chemical techniques have been applied to achieve CuI thin films which includes laser assisted molecular beam deposition [18], iodination of Cu thin films [19], pulsed laser deposition and electrodeposition [20], reactive sputtering [21], chemical bath method [22], successive ionic layer adsorption and reaction (SILAR) [17]. Here, we report a simple and efficient solid iodination method for the fabrication of γ-CuI thin films from reactively sputtered Cu_3N thin films. The chemical reaction of sputtered copper nitride (Cu_3N) with solid phase iodine yields a p-type γ-CuI thin film with high hole mobility and specific capacitance.

2. Experimental section

2.1. Thin film fabrication of Cu_3N

Cu_3N thin films were prepared by reactive RF magnetron sputtering system (Plassys MP 300) using a 50 mm diameter and 3 mm

* Corresponding author at: Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai 600119, India.
E-mail address: shyjusalammal1983@gmail.com (S. Thankaraj Salammal).

https://doi.org/10.1016/j.matpr.2019.05.387
2214-7853/© 2019 Elsevier Ltd.
Peer-review under responsibility of the scientific committee of the Advanced Materials for Clean Energy and Health Applications (AMCEHA).

Please cite this article as: K. k. Chinnakutti, V. Panneerselvam, D. Govindarajan et al., Fabrication of p-type cubic γ-CuI by solid iodination process for energy conversion and storage applications, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2019.05.387

CHAPTER 15

Synthesis, design, and characterization of controlled nanoparticles for the construction industry

Vinita Vishwakarma
Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, India

Contents

15.1 Introduction 323
15.1.1 Impact of nanoparticles and their significance in the construction industry 324
15.2 Synthesis and design of controlled nanoparticles and engineered nanomaterials 327
15.2.1 Physical methods 327
15.2.2 Chemical methods 329
15.2.3 Biological methods 329
15.3 Characterization of nanoparticles 330
15.4 Effect of nanoparticles on human health and safe handling 334
15.5 Environmental effects of nanoparticles used in construction industries 336
15.6 Discussion 336
15.7 Conclusion 337
References 338

15.1 Introduction

There are many challenges in the construction industry with respect to performance, efficiency, and environmental safety issues which need to be addressed. The developments in nanotechnology show significant promise in improving the traditional pattern of materials. The construction industry is going to benefit most from the implementation of nanotechnology. The conventional concrete properties will be improved with novel, smart, eco-friendly construction materials by using nanotechnology. The

Handbook of Nanomaterials for Manufacturing Applications
DOI: https://doi.org/10.1016/B978-0-12-821381-0.00015-6
© 2020 Elsevier Inc.
All rights reserved.



Plant extracts acting as corrosion inhibitor of steel reinforcement

Steel Reinforcement, Corrosion, Inhibition efficiency, Toxicity and Plant extracts

LAP Lambert Academic Publishing (2020-06-17)

€ 39,90

Buy at the MoreBooks! Shop

Corrosion of steel reinforcement in the construction field poses safety, financial and environmental threat. It can be handled effectively using inhibitors. Despite having an advantage, the synthetically produced inorganic and inhibitors impart huge cost and toxicity. As a result of extensive research toward green technology quite significant numbers of green corrosion inhibitors from plants shows promising results at low cost and eco-friendly option. The constituents in plants acts in a similar way like its chemical counterpart and provide best inhibition efficiency. This paper provides an overview of the types of corrosion, corrosion process, and mainly recent work done on the application of natural plant extracts as corrosion inhibitors for steel.

Book Details:

ISBN-13:	978-620-2-67049-4
ISBN-10:	6202670495
EAN:	9786202670494
Book language:	English
By (author) :	Annadurai Alagesan Gopikrishnan Venugopal Radhakrishnan Manickkam
Number of pages:	52
Published on:	2020-06-17
Category:	Technology

Distribution and Bioprospecting Potential of Actinobacteria from Indian Mangrove Ecosystems

11

Radhakrishnan Manikkam, Premalata Pati, Shanmugasundaram Thangavel, Gopikrishnan Venugopal, Jerrine Joseph, Balagurunathan Ramasamy, and Syed G. Dastager

Abstract

Actinobacteria are the ecologically and economically important prokaryotes with the unprecedented ability to produce novel metabolites. They are widely distributed in nature due to their physiological and metabolic versatility. Actinobacteria from marine ecosystem are recognized as a promising source for novel bioactive metabolites. Notably, mangroves are the most fertile and productive marine ecosystem due to their ecophysiological conditions and vast microbial diversity. At the global level, there are many novel actinobacteria which have been reported from mangrove ecosystems and also explored for their bioprospecting potential. However, in India, microbiologists have explored actinobacteria from selected mangrove ecosystems for the past five decades with special reference to culturable diversity and antimicrobial, antimycobacterial, anticancer, and enzymatic activities. So far, around 22 rare actinobacterial genera have been reported from the Indian mangrove ecosystems. However, several other novel actinobacteria with promising bioprospecting potential are yet to be explored. Hence, research on actinobacteria in the Indian mangrove ecosystem is still in its infancy. This chapter will describe the exploratory research work carried out on the importance, diversity, antimicrobial and

R. Manikkam · P. Pati · G. Venugopal · J. Joseph
Centre for Drug Discovery and Development, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, Tamil Nadu, India

S. Thangavel
DRDO-BU Centre for Life Sciences, Bharathiar University Campus, Coimbatore, Tamil Nadu, India

B. Ramasamy
Department of Microbiology, Periyar University, Salem, Tamil Nadu, India

S. G. Dastager (✉)
National Collection of Industrial Microorganisms (NCIM), CSIR-National Chemical Laboratory, Pune, Maharashtra, India

© Springer Nature Singapore Pte Ltd. 2019

T. Satyanarayana et al. (eds.), *Microbial Diversity in Ecosystem Sustainability and Biotechnological Applications*, https://doi.org/10.1007/978-981-13-8315-1_11

CHAPTER

10

Microbial leaching of heavy metals from e-waste: opportunities and challenges

V. Gopikrishnan[1], A. Vignesh[1], M. Radhakrishnan[1], Jerrine Joseph[1], T. Shanmugasundaram[2], Mukesh Doble[3], R. Balagurunathan[4]

[1]Centre for Drug Discovery and Development, Col. Dr. Jeppiaar Research Park, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, TamilNadu, India; [2]DRDO-BU Centre for Life Sciences, Bharathiar University Campus, Coimbatore, TamilNadu, India; [3]Department of Biotechnology, Indian Institute of Technology-Madras, Chennai, TamilNadu, India; [4]Department of Microbiology, Periyar University, Salem, TamilNadu, India

1. Introduction

Rapid urbanization, industrialization, and population growth have resulted in an increased demand of electronic and electrical gadgets. With fast technologic advancement and development, with the demand of people in parallel, industries are now manufacturing novel, superior, and smart electronic and electrical equipment (EEEs) at an alarming rate. More and more production, usage, and discarding of EEE products resulting in the generation of a huge amount of waste electronic and electrical equipment (WEEEs) (Tansel, 2017). The electronic waste (e-waste) normally consists of small and large equipment such as light emitting diode lamps, cell phones, smart mobile phones, superior televisions, refrigerators, printers, driers, temperature exchangers, as well as advanced computing devices (Bhattacharya and Khare, 2016). Due to the fast advancement in technology and tremendous market growth, most of these electronic and electric material have a very short life span, which results in the fastest accumulation in e-waste (Dave et al., 2017).

E-waste contains both hazardous and nonhazardous substances in its components. Globally, the annual e-waste generation is estimated as 20–50 million tons. The presence of organic toxic and hazardous substances in e-waste separates it from normal municipal waste. These hazardous substances like plastic, lead, mercury, cadmium, arsenic, etc., pose health hazards to humans when treated in uncontrolled condition via air, water, and soil.

The presence of these pollutants in the atmosphere creates various kinds of diseases. The people engaged in the recycling and recovery from the e-waste are severely affected with chronic and acute diseases like cancer. The infrastructure to deal with abundant e-waste in developing countries is insufficient in terms of technology, and more than 90% e-waste is treated with rudimentary and primitive techniques adopted by informal agencies. To prevent humans and the environment from contamination of these pollutants, it is desirable to study the characteristics of various hazardous materials in e-waste, so people are made aware about ill effects of e-waste treated in uncontrolled conditions (Vats and Singh, 2014). Various hydrometallurgical (the leaching of metals with acids or bases) and pyrometallurgical (the leaching of metals using thermal treatment) approaches have been reported for efficient metal recovery from WEEE; still a decisive assessment of conventional approaches has not been sufficiently explicated. Bioleaching is a low-cost, green technology for leaching metals from a variety of minerals and waste materials (Zeng et al., 2016). Bioleaching has several advantages compared to conventional hydrometallurgy or pyrometallurgy that require a large amount of chemicals with the production of high levels of environmental pollution (Zeng et al., 2013). Currently, a significant portion of many minerals/ores are processed at industrial scale using bioleaching (Walting, 2015). For example, approximately 20%–25% of the world's total copper (Cu) is produced using bioleaching (Brierley, 2008). A number of successful commercial bioleaching operations exist worldwide, and a Cu

Biovalorisation of Wastes to Renewable Chemicals and Biofuels
https://doi.org/10.1016/B978-0-12-817951-2.00010-9

Copyright © 2020 Elsevier Inc. All rights reserved.

Women Self Protecting System Using Internet of Things

M. Kavitha
Department of Electrical and Electronics Engineering
Sathyabama Institute of Science and Technology
Chennai, India
kaveem@gmail.com

V. Sivachidambaranathan
Department of Electrical and Electronics Engineering
Sathyabama Institute of Science and Technology
Chennai, India
sivachidambaram_eee@yahoo.com

***Abstract*— These days the safety of an individual is at stake, it may be due to ill health or due to the increasing crimes such as the sexual assaults, molestation, abuse etc. So in order to prevent these to a certain extent, this paper proposes an automated wearable smart device to prevent the above mentioned cause, which has access to internet (IOT). There are few bio sensors to sense the user's bodily changes and alert the required help when any abnormality is found as per the pre program of the device. The GSM and GPS are used to identify the victim's location when in need. The IP address of the Arduino is linked with the web server. The victim location is shared to the near by police station and to the pre registered mobile. While the buzzer alerts the surroundings of the victim.**

Keywords—IOT, Biosensors, Arduino

I. INTRODUCTION

In the modernized lifestyle with decrease in the ethics and morals, there is a rapid increase of crimes. The technology which has a major role in our day to day lives has both its own advantages as well as disadvantages, and when used correctly can make a great difference. One's safety is always at stake let it is due to sudden illness or due to sexual violation. These can be prevented by a new kind of technological development and its usage. Few of the recent technologies include mobile or android applications for ones safety such as Anuti, Wosapp, Reach-360 etc. These apps require the registration and login certain credentials , during the time of need the user has to wave the mobile for the message alert to be sent to the registered family and the authorities numbers for help[2]-[4]. The main drawback is that there are rural areas where people don't use mobiles, and in other cases where there are children, they often play with mobiles and if the accidentally happen to wave the mobile then it results in false alarming. Some other technologies prescribe wearable accessories which need manual operation i.e. manually whenever there is need of help they need to press a button to alert for help using various communication technologies such as RFID, Zigbee, Raspberry pi, Bluetooth etc[5]- [11], but these cannot be that effective as sometimes the victim might be in the condition to not able to press the button. The other way is where the device is clipped to the user's foot and there is a certain gesture to be made for activation of the device [12] which might not be possible when danger comes. While the other suggest to have the device in the public transportation and to sign in the using One Time Password [13] but this is limited to public transportation. Other suggests a shock generating glove [14] which is not a mandatory accessory in few countries or except for only winter. The most updated are wearable with automation using IOT [15]-[19]. All the technologies discussed are only for women safety, but whereas we propose for all genders, ages and both health and prevention of any sexual violation.

II. EXISTING SYSTEM FOR WOMEN SAFETY

Fig. 1 represents the block diagram of one of the recent existing system for women safety [1]. According to the block diagram, the system is a smart band which consists of a microcontroller embedded with a few bio sensors such as motion sensor, temperature sensor and pulse rate sensor, a BLE module and a power supply. This smart band is connected to the mobile through the Bluetooth module. The sensors frequently monitor the body condition of the user and when an abnormality occurs within the body, which is pre programmed in the band compares and when abnormality is Confirmed then the collected data is acquired and activates the pre installed app which in return sends the messages to alert the family, police and the people who use the designed app.

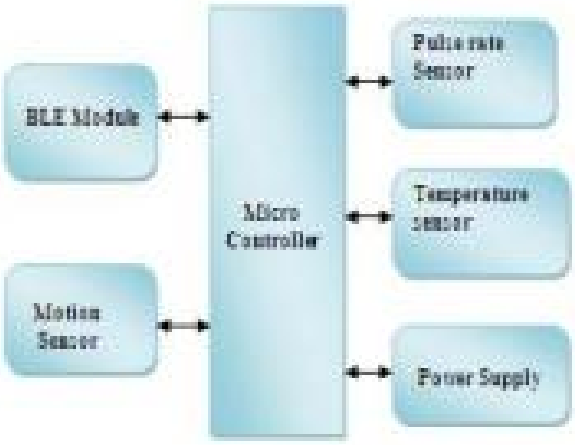


Fig. 1. Block diagram of the Existing System

978-1-5386-1508-9/18/$31.00 ©2018 IEEE

Proceedings of the International Conference on Communication and Electronics Systems (ICCES 2018)
IEEE Xplore Part Number CFP18AWO-ART; ISBN: 978-1-5386-4765-3

A Women Safety Portable Hidden camera detector and jammer

Sindhu.K, Dr. R. Subhashini, Dr.S.Gowri, J.S.Vimali
Department of IT, School of Computing,
Sathyabama Institute of Science and Technology
Chennai-600119, India.
Sindhukb9@gmail.com, Subhashini.it@sathyabama.ac.in, gowri.it@sathyabama.ac.in, vimali.it@sathayabama.ac.in

***Abstract*— In recent times our country had developed in many ways but still many crimes against women exist. There are no proper means to detect the hidden camera which now had become a serious issue. This device intends to maintain the safety and security of people by detecting the spy camera and stop the camera working. Since hidden cameras are less cost and small in size, it became easy to install camera almost everywhere. It finds its application in the areas where actually cameras are not allowed. This device works with a principle that camera operates in the frequency range 0.1GHz to 3GHz, so RF signal detector detects the frequency range and if it is within this frequency range it alerts us with an alert light. A video camera can also be detected by this detector by extending the detecting range from 0.05 GHz to 3 GHz we can detect both camera and video camera. Once, it is detected it alerts us with a LED light and it an option to stop the camera transmission by creating strong interference between the camera and base station. This device also comes along with an eyeglass which can be used if alternating pulsating frequency or if recording camera is used where RF signal detector cannot be used. Cameras can be detected by flashing red light once viewed through optical glass viewport. The flashing light will reflect off the 'spy' camera's lens and will be able to spot it.**

Keywords*— *RF signal detector, security, frequency range, jammer.

I. Introduction

The issue of hidden cameras at public places became a severe threat to people these days. The cameras are secretly put up in the changing room, theatres, restrooms and other public places which affect the privacy of the people. Some people might argue that cameras are easy to find and this is unnecessary, but this is really a tough job to find a hidden camera in a room. Most of the camera is wireless and hence became very easy to hide. Manually, finding a camera is almost impossible. This is based on the fact that almost all the camera devices work within the frequency range 0.1GHz to 3 GHz. By this, we can prohibit the unauthorized use of the camera by using a detector that senses the presence of hidden camera as they radiate RF signal for transmission. By this, we can even detect the presence of small video cameras whose frequency range is 0.05GHz to 3GHz (which is a quite difficult task as we need to differentiate the video signals from WIFI signal and TV signal). The detector detects the RF signal and the signal information is indicated by the LED light and provides us an option to stop the camera data transmission by generating the same type of interference so that the receiver or base station will not understand which signal to receive. This also has an eyeglass which can be used to detect the camera hidden with alternating pulsating frequency or if recording camera is used instead of a camera which uses Wi-Fi to transmit information. The principle on which this hidden spy camera detector works is alternating pulsating laser frequency using which hidden cameras can be detected as flashing red light once viewed through optical glass viewport.

II. Existing System

In the existing system, there are many hidden cameras detecting mobile application. They use smartphone flash to light up camera lenses and detect them. Some generate electromagnetic signals and with the help of the interrupt in the signals, they detect the presence of a hidden camera. The advantage is cost and reliability. But this is cover only the smaller distance, and it works properly only if the mobile has the magnetic sensor. Moreover, it should be done manually (i.e.,) the checking should be done by moving the mobile closer to all the places in the specified area. In addition, there is no video camera detector as it is hard since it should be differentiated from TV signals & WIFI signals. The proposed device can cover a comparatively large distance. It also provides an option to jam the camera network to stop the transmission. The proposed system can also detect the video camera which can be hidden smoke detector.

III. Proposed System

The proposed device will have RF signal detector and jammer, the moment the detector detects the RF transmission signal from an activated camera, it starts to alert us with a LED light which will be on the device. The LED continues to glow until

Wireless Pers Commun
DOI 10.1007/s11277-017-4948-y

Wireless Soil Monitoring Sensor for Sprinkler Irrigation Automation System

G. Nagarajan[1] · **R. I. Minu**[2]

© Springer Science+Business Media, LLC 2017

Abstract In a developing country like India, there is an exponential rise in population nutrition requirement. To meet up with both the ends, the agricultural techniques should be perfected for optimal yield and quality. Irrigation and soil property monitoring system using sensors can be automated and operated wirelessly to achieve optimal water supply control and surveillance. The objective of this paper, is to automate the whole wireless sensor network (WSN) system with a control over water pumps and dripper valves. The humidity, temperature and pH sensor's percepts provide a feedback, to control the water content of the soil. The system has an low-cost and energy reliable ZigBee for sensor data transformation, high-range GPRS system for data storing and analysis, and the whole system is powered by Solar panels which makes it self-sustainable. Customizable options for different crop with different requirements make it a versatile WSN system for automated irrigation based water management.

Keywords Agriculture · Irrigation · Microcontroller · Soil moisture · Soil pH · WSN · ZigBee · Solar power

1 Introduction

Agriculture in India is still one of the largest sectors to contributing 17.9% General Directorate of Pharmaceutical Affairs (GDPA) to the economy till 2014. Over 58% of the rural households depend on agriculture as their principal means of livelihood. While such a

✉ G. Nagarajan
nagarajanme@yahoo.co.in

R. I. Minu
minu@jerusalemengg.ac.in

[1] EEE Department, Sathyabama University, Chennai, Tamil Nadu, India

[2] CSE Department, Jerusalem College of Engineering, Chennai, Tamil Nadu, India

Published online: 09 September 2017

Wireless Data Acquisition System for Underwater Environment

K.H.Shakthi Murugan., M.Sumathi
School of Electrical and Electronics Engineering, Sathyabama University,
Chennai-119, Tamilnadu, India
shakthimuruganeee@gmail.com , sumagopi206@gmail.com

***Abstract*—Underwater devices are gaining importance in today's environment as it can collect data that is difficult for human beings to gather. It mainly deals with the study of marine environment and to check the purity content of the water. The main objective is to design an underwater robot for monitoring purpose and to have a better understanding of water and the marine environment. This paper deals with the implementation of data acquisition system using MSP430 microcontroller and Anaren Transceiver. It is an efficient and effective wireless network that can acquire data from water wirelessly and transmit the same to the computer. pH, temperature and CO_2 sensors acquires the data and transmits it to the Anaren chip. Here MSP430 controller acts as the central data acquisition system and is used for controlling the overall process. This method of implementation will be useful for cleaning, monitoring and understanding of the underwater environment. The results obtained display the values of temperature, gas and pH for each sample taken. The sample taken can be found if acidic or basic. As an extension, the nature and turbidity of water is also detected using pH sensor and LDR sensor.**

Keywords—data acquisition; underwater; anaren; pathogens; sensors

I. INTRODUCTION

Nearly 75% of the surface of earth is covered with water. Water is an important resource with any life and is available in the form of oceans and seas, rivers and lakes, ponds, etc. Every person needs a certain amount of water for their activities: drinking, bathing, washing hands, washing your clothes or your house. However, there are people who wasted no thought that only a small part can be harnessed for human activities. The average cost of water varies from region to region and depends on the activities that everyone performs. Many new technologies have been proposed to explore this vast aquatic environment. These technologies provide us with a number of scientific and environmental applications. Some of the applications are pollution monitoring, oil/gas spills monitoring, disaster prevention in terms of tsunami warnings, navigation and surveillance. Therefore, this unique feature of underwater communications has created an interest and many research activities are in progress for underwater networks. Data acquisition is the process of sampling signals that measure real world physical conditions and converting the resulting samples into digital numeric values that can be manipulated by a computer. Data Acquisition Systems, abbreviated by the acronyms DAS or DAQ, typically convert analog waveforms into digital values for processing. The components of data acquisition systems include: Sensors, to convert physical parameters to electrical signals; Signal conditioning circuitry, to convert sensor signals into a form that can be converted into digital values; Analog to digital converters to convert conditioned sensor signals to digital values. There are also open-source software packages providing all the necessary tools to acquire data from different hardware equipment. These tools come from the scientific community where complex experiment requires fast, flexible and adaptable software. Robotic Underwater Vehicles (RUV) is used in underwater aquatic bodies to study their underwater environmental characteristics [1]. Many attributes of the underwater environment interfere with the detection, characterization, and recovery of military munitions. Except at very shallow sites, munitions under water are difficult to access. Conditions interfere with the ability of sensors to detect and characterize them and remediation is more difficult. The underwater environment has a lot of research challenges yet to be explored [2]. The cost of commercial data acquisition systems is a great challenge, as they are costlier [3]. New technologies for efficient data transfer between underwater wireless communications have emerged [4]. Underwater communication has been made possible by exploring technologies like acoustics [5, 6].

Water is a need for everyone and it is getting contaminated due to pathogens. The quality of water used by human beings is a global concern [7, 8]. While most microbes play an important role in nature, certain microbes that are potentially harmful can contaminate water, and cause a plethora of infectious diseases in both animals and humans. Impure water causes illness leading to different diseases and is even fatal in some cases [9, 10]. The World Health Organisation has revealed that 2.2 million people around the world are dying due to consumption of contaminated water annually. Nearly twelve billion US dollars economic loss is caused globally every year by using low quality water [11]. Thus improving water quality is a major concern which will reduce the number of people affected by waterborne disease every year [12]. Unfortunately, there are a number of obstacles in achieving this target and many methods are proposed [13]. Some specific species of pathogens can be detected by using molecular methods [14, 15&16]. Drinking water has an alkaline pH between 7.0 and 9.5 indicating that water is rich in minerals such as calcium and magnesium. Water with a pH of 7 or above and an electron activity in the negative milli-volts with a redox of 22 or less is considered the best water to drink.

978-1-5090-3239-6/17/$31.00©2017 IEEE

Vehicle Quality Check Test Bench Using Arduino

PABIRAMI
Research Scholar, EEE
Sathyabama
Institute of science and technology, Chennai
Abiramiramkumar80@gmail.com

M.PUSHPAVALLI
Research Scholar, EEE
Sathyabama
Institute of science and technology, Chennai
Pushpa.murugan@gmail.com

P.SIVAGAMI
Research Scholar, EEE
Sathyabama
Institute of Science and technology, Chennai
sivaganitec@gmail.com

S.SINDHUJA
Research Scholar, EIE
Sathyabama
Institute of Science and technology, Chennai
aeonian4u@gmail.com

***Abstract* – In our present scenario, so many road accidents are happening because of poor quality of surface of the road. To avoid accidents and to provide safety for the driver and passenger the subsystems of the vehicle like Engine firing order, Exhaust gas, Anti-lock Braking System, Vibration should be monitored and controlled continuously. Wireless Sensor Networks (WSNs) are applicable in numerous domains, including industrial automation where WSNs may be used for monitoring and control of vehicle and equipment. However, the requirements in the vehicle test bench differ from the general WSN requirements. In recent years, standards have been defined by several vehicle monitoring alliances. In this proposed work the subsystems of the vehicle are monitored and informed automatically to the driver through arduino UNO board. Thus this automation reduces the human error caused due to manual monitoring and mostly the accidents can be avoided. Also in this paper anti lock braking system is discussed to avoid skidding of the vehicle during sudden slowing mechanism is applied.**

Keywords - Anti-lock Braking system, Wireless Sensor Network, arduino uno board.

I. INTRODUCTION

In this paper, the various vehicle subsystems that are prone to be accidental are checked using IOT. The driver gets information about the status of the vehicle through mobile or any communicating devices. By this proactive information majority of the accidents can be avoided.

The information regarding the vehicle condition is obtained through various sensors connected with arduino board. Arduino is a microcontroller based board and is one of the advanced IOT technologies for automation. The different sensors used in this work are fire sensor, vibration sensor and gas sensor. Erkin Dinemen explains about the ABS control extremum-seeking algorithm depends upon the slip ratio between the tire patch and the road [1]. Hossein Mirzaeinejad describes the safety of vehicles in anti-lock braking system obtained by regulating the wheel slip at its optimum value [2].

II. PROPOSED SYSTEM

Fig 2.1 represents the block diagram of the proposed system. The block diagram consists of arduino board and connected sensors. In proposed system engine firing order timing, exhaust gas quality, ABS brake testing, communication module testing are verified using this test bench. Exhaust gas quality is checked using MQ gas detectors to check the quality of exhaust. This is essential to check the quality of exhaust gas because under operating condition the CO gas flow is maintained uniformly without causing any damage to the persons inside the vehicle. Suppose if the engine is stopped we have to take care of the exhaust gas quality to avoid death due to leakage of poisonous gas from the ventilating system [3]. For engine firing order at each firing of the ECU (led) the controller checks the analog input of the sound sensor (Push button) to detect the exact firing to check engine efficiency.

There exists an engine system and battery system in a vehicle. The causes of fire in a vehicle may be due to high temperature, leaking fluids and abnormal wear of the engine and battery [4]. ABS brake testing is done by enabling the ABS lock(led), enabling the drive mode in the transmission (DRIVE MODE LED) and detects any movement of wheels by using the ABS encoder movement (analog pin monitoring) and finally signal test is done by sending a test message via vehicle communication section (tx zigbee) and waits for its reception within a particular time window in receiver side (rx zigbee) [5-9]. All these system are emulated to show the working of the test bench.

In chapter I introduction about the vehicle subsystems and the methods for preventing gas leakage, firing, Vibration and ABS system is discussed. Chapter II deals with the detailed description of the block diagram of proposed system. The existing system is of manual controlled vehicle quality check test bench. Chapter II dealt with the automatically controlled vehicle quality check test bench by using Arduino UNO board. Chapter III dealt with the simulation results of the proposed system unit for various sensors activation level. The proteus system is used to simulate the proposed work and the coding is written using arduino coding. Chapter IV gives the detailing of hardware implementation and hardware results. MQ-X Sensor is used as the gas sensor to detect the leakage of exhaust gases, Infra red flame sensor is used to detect the firing engine order of the vehicle, a vibration sensor is used to detect the vibrations of the vehicle due to potholes, speed breakers etc., and a ABS enabling system is activated using ABS lock activated by using Zigbee transmitter and receiver units. Followed by this chapter V concludes the proposed work.

978-1-5386-1508-9/18/$31.00 ©2018 IEEE

Indian Journal of Science and Technology, Vol 9(5), DOI: 10.17485/ijst/2016/v9i5/87201, February 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Urbanization and its impact on Urban Heat Island Intensity in Chennai Metropolitan Area, India

Lilly Rose Amirtham*

Sathyabama University, Chennai – 600119, Tamil Nadu, India; lillyini@gmail.com

Abstract

Background: Urbanization results in the formation of heat pockets termed as "Urban Heat Island" (UHI) and their intensities depend on urban built geometry. This study attempts to assess the impact of urbanization on urban heat island intensity in Chennai Metropolitan Area (CMA). **Methods/Analysis:** Air temperature and relative humidity were measured at one hour interval at 30 fixed locations covering urban, suburban and rural locations in and around CMA. HOBO data logger (H08O U10 Temp/RH), enclosed within a white perforated plastic box to shield the instrument from direct solar radiation were deployed for the study. UHI intensity was computed based on the reference data from Numgambakkam Meteorological station. Day time and night time temperature isopleths of CMA were derived using ArcGIS, including the isopleths at maximum and minimum temperature occurrences at 14.00 hrs and 06.00 hrs respectively. The temperature isopleths at 10:00 hrs, 14:00 hrs and 18:00 hrs were analyzed for daytime intensity of UHI and their temperature distribution pattern. **Findings:** Maximum daytime temperatures were recorded between 11:00 hrs and 15:00 hrs which varied between locations based on urban morphology and the time at which the measurement location is exposed to incident solar radiation. The daytime temperature isopleths at 14.00 hrs revealed the existence of a cool island in the urban core of CMA with a temperature difference of 10.4°C in summer and 3.7°C in winter. The night time UHI in the CMA was analysed with the temperature isopleths of 22:00 hrs, 02:00hrs and 06:00hrs. The night time temperature isopleths reveal the significant existence of positive UHI in the CMA. The isotherms at 06.00 hrs revealed a significant positive UHI ranging from 3.6°C in summer to 4.1°C in winter. The results of the study indicate a maximum nocturnal UHI during the calm, clear, winter periods in the CMA. The urban rural differences also revealed a significant cool island during the day with the maximum cool island intensity during summer. **Applications/Improvements:** A significant correlation between urban-rural differences and density of urban built form was established; indicating the significance of building regulations that define urban built geometry, in designing comfortable urban environments.

Keywords: Built Geometry, Cool Islands, Chennai, Urbanization, Urban Heat Island Intensity

1. Introduction

Urban areas experience a distinct climate due to the transformation of natural ground cover into hard impervious land parcels. When compared to rural surroundings, urban activities associated with increased anthropogenic emission, increases air temperatures significantly. The urban areas experience higher temperature than the rural areas as the built fabric stores the absorbed incident solar radiation and anthropogenic heat released from vehicles and equipments[1], resulting in the formation of heat pockets that are termed as "Urban Heat Island" (UHI)[2]. And the intensity of heat island is determined by the temperature difference between urban and rural areas. The intensity of heat islands depends on the density, population and size of a city[3] and the morphology of cities[4]. The urban heat island measured through the differences in ambient air temperature at the canyon layer (between buildings at street level) is termed as the Canyon Layer Urban Heat Island. Urban geometry and capacity of urban fabric to absorb heat determines the intensity of UHI predominantly[5].

*Author for correspondence

Unsupervised Spectral Sparse Regression Feature Selection using Social Media Datasets

R. Sathya Bama Krishna[1]
[1]Research Scholar, Faculty of computing, Sathyabama University, Chennai, India.
rsathyarajeswari@gmail.com

Dr. M. Aramudhan[2]
[2]Professor and Head, Department of Information Technology, Perunthalaivar Kamarajar Institute of Engineering and Technology, Karaikal, India.
aranagai@yahoo.co.in

ABSTRACT
Recent innovations in the social media sites such as Google Plus, Face book, Thomas cook, Twitter and Flickr have increased more and more subscribers to participate in social online events just as connecting with friends, families and other like-minded subscribers, posting statuses and blogs, sharing and communicating information etc. As a result usage of social media by the users has increased and generated unprecedented volume of social data such as linked, labeled and unlabeled data. Mining the data produced by the social media has its ability to dig out actionable patterns that could be valuable for government, business, users, and consumers. However, Social media documents are vast, dynamic, high dimensional, unstructured and forcible thereby open to new challenges. At present, in data mining Supervised Feature Selection method outperforms effectively in manipulating labeled and high dimensional data thereby providing an effectual mechanism for learning. However, there is no single method proven to be effective for mining the high volume of data produced by social media. An Unsupervised Feature Selection method employing spectral analysis taking social media dataset have been proposed. This paper also aims at conducting experiments with the data sets from different social media to evaluate its accuracy and performance.

Keywords
Data mining; spectral analysis; social media; actionable patterns

1. INTRODUCTION
In general, choosing the subset of similar and relevant features is known as the process of feature selection. This technique of feature selection involves the data that includes both the redundant and irrelevant feature instances [7]. No more information than the currently selected features can be provided by the redundant features and no useful information is provided by the irrelevant features in any of the context.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Request permissions from Permissions@acm.org.
ICIA-16, August 25-26, 2016, Pondicherry, India
© 2016 ACM. ISBN 978-1-4503-4756-3/16/08…$15.00
DOI: http://dx.doi.org/10.1145/2980258.2980323

Supervised feature selection methods are used to retrieve relevant labeled feature set from the authentic features that aid classification, clustering and retrieval [9]. The feature selection problem is importantly an optimization problem which is expensive computationally. Classical feature extraction techniques had addressed these drawbacks by choosing the topmost features upon the basis of specific scores that are independently measured for all given features. But this approach does not care for the possible correlation of different features and hence it's not able to retrieve an optimized feature subset. This provides an inspiration to develop the unsupervised feature selection [11]. The unsupervised feature selection is based on machine learning technique called unsupervised learning. The unsupervised learning method finds out the hidden structure of unlabelled data. The hidden structure could be constructed by training with manually extracted patterns from the given data [8].

2. RELATED WORK
In this section, let us overview the feature selection for linked data and community detection algorithms, and the recent technologies emerged.

2.1 Feature Selection for Linked Data
Social media has large volumes of unlabelled high-dimensional data. Feature selection mechanism plays a key role to deal with those data. The relevant feature are obtained by social correlation theories like homophily and social influence where homophily states that the users with similar interests are likely more linked and social influence states that linked users are more often tend to have same interests [4]. These theories are used to extensively evaluate the four types of relations namely co-post, co-following, co-followed and following. But it does not fit well to any other relations such as group of users. Another drawback is that some relevant information is hidden as social media is inherently linked and also it does not uses spectral way for selecting relevant features.

2.2 Detecting Social Media Community
It is done for online forum sites where users hold conversations in the form of messages. It focuses on the dynamic interactions of the users and detection of communities. Users having the similar interests are identified using the related events posted by them. Sometimes these users might not interact with each other by their events alone though they have the similar interests. So the temporal coherence analysis algorithm is proposed to find out the temporal similarities between the users which are the input for the soft community detection algorithm that works on the principle of soft clustering to retrieve the user of one or multiple communities. It is done with the help of spectral analysis that evaluates the dynamic interactions between the users where they are denoted by

Two-Dimensional Stockwell Transform Based Image Fusion for Combining Multifocal Images

Ella Madhava Babu
Final Year Student,
Dept. of ECE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

S. Dusyanth Maniks
Final Year Student,
Dept. of ECE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

**N.M. Nandhitha
Prof. & Dean,
School of EEE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

N.Selvarasu
Prof. & Head,
Dept. of ECE,
SVS Group of Institutions,
Warangal,
Telungana.

S. Emalda Roslin
Associate Prof.,
School of EEE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

Rekha Chakravarthi
Associate Prof.,
School of EEE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

M. S. Sangeetha
Assistant Prof.,
School of EEE,
Sathyabama University,
Jeppiaar nagar,
Chennai 600 119.

**Corresponding Author: nandhi_n_m@yahoo.co.in

***Abstract*- In real time image acquisition, especially in the areas of remote sensing and telemedicine, capturing the entire scenario/specimen/scene is not possible even with high resolution sensors. Hence repeated image analysis is done on a large set of images. It increases the computational complexity of the analysis system especially when a large set of images is to be analyzed. Hence it is necessary to identify a technique that reduces the computational time. One such technique is image fusion which allows a set of images to be combined. Image fusion is defined as the process of combining two images which not only retains complexity info but also strengths redundant information. Conventionally wavelet transforms are used or decomposing the images into low and high frequency components. Later choose maximum and choose average are used for combining the coefficients. Though wavelet transfer provides the spectral information, it could not provide details about the phase or phase difference. Also in case of bursts in images, wavelet transform could not provide the appropriate information. Hence it is necessary to identify a suitable transform for obtaining the burst details. Stockwell transform provides frequency, time and phase information. In this work, a set of multi focal images are acquired and Stockwell transform is applied on this image. Later appropriate fusion rules are used to combine this features and inverse Stockwell is used on the fused image to obtain the original image. Performance of the proposed technique is measured in terms of Root Mean Square Error and standard deviation. Performance of the proposed technique is compared with that of wavelets based image fusion and is found that Stockwell transform provides better results (both quantitatively and qualitatively) than wavelet transform.**

Keywords- Stockwell Transform, Image Fusion, Multifocal Images, Root Mean Square Error, Standard Deviation

I. INTRODUCTION

With the advent of technology, Exabyte's and petabytes of images are captured processed and transmitted every day. These images are processed to extract information, especially in video or in image processing where large number of frames are monotonically processed the computational complexity is the major issue ,the above said problem is inherent in telemedicine (processing problem is inherent),remote sensing. Hence in these applications it is necessary to reduce the computational complexity by reducing the number of images which are being processed. Image fusion is such a technique that combines images in order to provide both the redundant and complementary information that is the redundant information is strengthened ,while the complementary information is also present .Image fusion is broadly classified in to 2 types namely (1).Spectral image fusion (2).Temporal image fusion. In spectral image fusion as the size, focusing area of the sensor varies, image registration is necessary for example spectral image fusion is favored when multiple sensor image fusion is done. In temporal image fusion, image registration is not needed as all the images are of same size but only taken at different instances of For instance in video surveillance and MRI analysis a multitude images are taken at

Track Health Care Using Location Based QoS Web Service Recommendation System

Y. Justin Dhas
Research Scholar: Faculty of computing
Sathyabama University
Chennai, India
justindhasy@gmail.com

P. Jeyanthi
Professor: Faculty of computing
Sathyabama University
Chennai, India
jeyanthiprabhu@yahoo.com

***Abstract*— Web services are referred as an open standard (XML, SOAP, and HTTP etc.) based Web- applications that interact with other Web applications for the purpose of exchanging data. Collaborative recommender system recommends Items to the user based on the actions that other users in the same environment did in the past. Web service recommendation plays an important role for both the service providers and active users in building reliable service - oriented systems. It has become vital to resolve typical problems of recommendation system due to the increased number of Web services advertised on the Internet. Since the traditional service recommendations do not have much concern about the composite service so in this paper we use the QoS values to recommend the best service. This paper focuses on recommending healthcare, health specialists and diagnostic centers to the health-seeker. Thus for this purpose the health-seeker location is tracked via IP Address and mapped on the Google map to track down the target location. The list of health care, health specialists and diagnostic centre will be displayed based on the historic QoS values. Finally, every user has to rate the health specialists/ medical practitioners, health care and diagnostic centers to give the QoS values.**

Keywords—webservices ranking; recommendation system; collaborative filtering; service oriented architecture

I. INTRODUCTION

In recent years, we have witnessed that there is a tremendous increase of Web services on the Web. This understudy prompts Web service revelation which plays a noteworthy issue in benefit processing [1]. To recommend Web service, we employ collaborative filtering technique [2]. This collaborative filtering algorithm is mainly used in commercial recommendation systems [3]. There are 2 techniques by which collaborative filtering is categorized, memory-based approach and model-based approach [4]. Furthermore, depending on the user relationship memory based is classified into two types: 1. User-based approach [4] 2. Item-based approach [5] User-based approach is used to recommend particular items to the user by collecting previous user values. Item-based approach is used to recommend the items to active users that are similar to those items users preferred in the past.

In order to recommend the Web service with non-functional properties, quality of service (QoS) is used [6]. QoS mainly deals with predicting missing values for the target service. Thus to predict the Qos, collaborative filtering algorithm [2] is used. QoS mainly includes few performance factors like reaction time, cost, quality etc. Thus along with this, in our paper we use location based QoS prediction approach using collaborative filtering algorithm to recommend best health cares and diagnostic centre's to the health-seekers. Thus along with this, in our paper we use location based QoS prediction approach using collaborative filtering algorithm to recommend best health care and diagnostic centre's to the health-seekers.

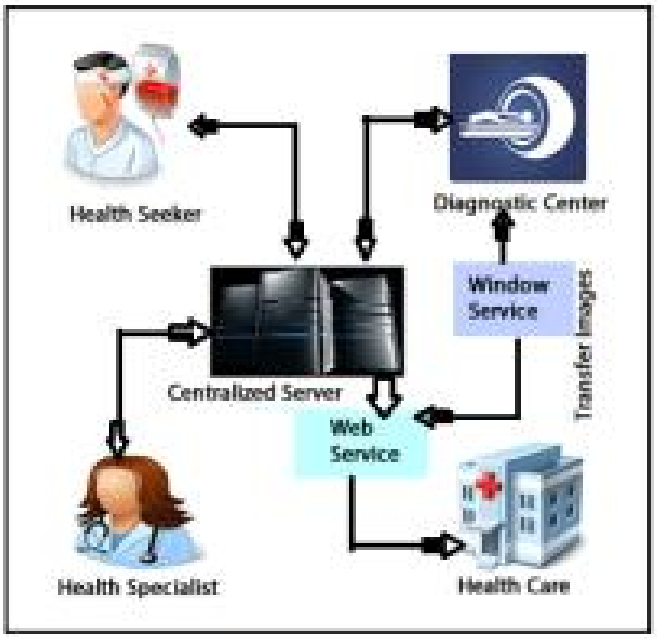


Fig. 1. Overview of recommendation system

Fig. 1. depicts the overview of recommendation system for health care, health specialists and diagnostic centre's. Initially a centralized server is maintained. The health seeker registers and logins into to website to view the recommended list. Similarly, the health care and the health specialists and the diagnostic centers need to be registered in the server. We use location-based QOS prediction approach to recommend a list of health care's or heath specialist or diagnostic centre's to the health-seeker. Then he selects from lists and other processes are carried out. After his treatment is completed, he has provided us with the feedback of the service during his treatment.

978-1-5386-4349-5/17/$31.00 ©2017 IEEE

This full-text paper was peer-reviewed and accepted to be presented at the IEEE WISPNET 2016 conference.

Text Separation in Document Images through Otsu's Method

M.Siva Sindhuri[1] ,N. Anusha[2]
Department of CSE,
Sathyabama University,
Chennai
sindhurisiva@gmail.com[1] ,anusha.nallapareddy@gmail.com[2]

***Abstract* - In the process of analyzing the document images, the text separation from the background plays a vital role. If the quality of document image is good, then text can be separated very easily by applying simple techniques of thresholding. Whereas, in noisy images it requires very accurate analysis. Various thresholding techniques that are available are efficient and produce better results, but they are very slow in that they require threshold calculation for each and every pixel. In this paper, we overcome this problem by proposing an Otsu's method in which it follows a recursive method for segmenting an image. The experimental results show that the proposed Otsu's approach is able to extract the text very effectively from its background even if the input is a noisy image.**

***Index Terms*- Document analysis, Otsu method, Image segmentation, Text separation.**

I. INTRODUCTION

Libraries, museums and major institutions have a variety of documents, which include manuscripts, civilizations, archives, historical documents etc. All these documents might degrade over a period of time. If so, it becomes very difficult to extract the text from the background of document images. The main cause of degradation is improper preservation. So, in order to preserve these documents, it needs to save as images. But the technique of digitization alone is not sufficient. This requires a lot of other techniques and tools for analysis. In the traditional methods of document analysis, image binarization technique is used. But this binarization technique is very critical in which it cause very bad separation and leads to loss the wanted information and may add unwanted information or noise [1]. And in today's life the text mining studies have very importance due to the rapid growth of electronic documents. For any type of text documents web is acting as the primary source. But this information is stored in unstructured format. So the need for retrieving of useful information from multiple sources of data requires the analysis of textual data [2].

A. Related Work

[1] Describes a new technique called hybrid thresholding. This technique is mainly performed in two steps, i. e, globally and locally. This method increases the speed and performance and it also extracts the text from degraded document images. But it can't produce better results in noisy images. [2] Concentrates on highlighting the various methodologies and techniques which are used for the classification of text documents. It also reviewed many methods on text mining and document classification. But this paper doesn't show any experimental results for image classification. [3]Proposed a new technique called novel document binarization method. In this technique, a contrast map is constructed for the degraded document image. This novel binarization method is very robust and simple. But this technique makes use of local image contrast in which it depends on local maximum and minimum.

[4] Proposed a segmentation algorithm which is based on Markov Random Field. Here the image is segmented into very small regions. And in this paper a split-and-merge procedure is also used to improve the accuracy of segmentation. The fast segmentation algorithm is very fast since it operates on smaller regions. But it becomes very complex when the tree grows. [5] Proposed a phase based binarization model. This model is mainly performed in three steps i.e, preprocessing, binarization and post processing. This method also generates a tool called ground truth generation to speed up the process. But this method doesn't suit for multispectral images. [6] Proposed an algorithm for segmentation of context-based visual data. Here the segmentation is performed by a probabilistic model. Here we can improve the segmentation quality by a semi supervised method. This algorithm considers only color features. Others like regions of texture, shape etc., are overlooked.

978-1-4673-9338-6/16/$31.00 ©2016 IEEE

Survey on Threats in Online Social Media

Soumya.T.R and S. Revathy

***Abstract*—In recent years Online Social Networks are raising day by day. People are attracted towards the use of Social Networks due to its virtual method of Socializing people. Young Generation are in large numbers who are currently involved a lot with the Social Networks. Along with the increasing usage of Online Social Networks, Security threats have been increased simultaneously. There are different types of threats currently prevails which make the users as victim for Security attacks called as Cyber Security Risks. In this study we will discuss about similar threats to networks and the proposed solutions to protect the networks from such threats.**

***Index Terms*—Threats, Online Social Networks, Cyber Security, Malicious, Vulnerability**

I. Introduction

ONLINE Social Networks are found as the integration of people in a virtual network. This platform is used to communicate with friends and family and is effective for people to keep in touch even in the geographic separation. Lots of sharing are performing day by day which may be of photos, videos, word contents and many forms of data based on the users and their interests.

Nowadays more amount of conversation are performed through Online Social Networks [1-3] and became an unavoidable part of current life. High amount of time have been spend by people especially youngsters.[2] Some of the popular Social Networks are Facebook, Tweeter, Instagram, YouTube,etc.[3-5] Apart from entertainment Social Networks are also used for the purpose of Education, Job Search, etc, [3]

Due to the existence of large amount of data in Social Networks, social networks are found to be a desirable source for Hackers whose intention to get the information of an individual or an organization without their knowledge. Since most of the information in the social networks are fresh and up to date, a lot details can be extracted directly as well as indirectly. For example, the posts liked by a person which may reveal the character or interests of that person[6-10]. Similarly, a lot more information can be obtained by the attackers or hackers. In this paper we will study different threats to online social networks and solutions to these threats[11-16]. The rest of this paper is organized as follows. Section II explained different threats in online social networks. Section III describes the solutions to eliminate security risks in social networks. And section IV is the conclusion.

Soumya.T.R is Research Scholar with the Sathyabama Institute of Science and Technology,India (e-mail: soumyatr.soumya@gmail.com).

S. Revathy is Asst.Professor with the Department of Information Technology, Sathyabama Institute of Science and Technology, India. (e-mail: ramesh.revathy@gmail.com)

II. Big Data

The term Big data means that large amount of data sets which cannot be normally deal with the traditional applications or software used for data processing. Several processes such as collecting data, storing the data, analyzing the data, searching data, share or transfer data, update the data, maintaining privacy of data, etc. Data can be classified as Structured, Semi-Structured, Un structured whereas Big data deals mainly with the unstructured data. When the growth of data is in massive scale as terabytes and peta bytes, Big data needs some techniques along with some technologies to focus on the complex and diversing datasets. The Big data, based on size, speed and types or sources of data , usually classified as 3Vs: Volume , Velocity and Variety. 'Volume' refers to the size, is the main reason to call these as big data. 'Velocity' means for the speed of growth and processing of data responsible for the development. 'Variety' is the nature of data which is responsible for the effective analysis of data. Big data is a large amount of information assets which needs a processing system which is of automated, cost-effective and with the ability of decision making .The growth of data analytics and data management field in industry which is growing much faster than the software business just because of big data. Big data is used in several services such as Finance, Manufacturing, Health care, Real estate, Media, Education, Research, Sports, etc. Advanced tools which can be used to retrieve information from the large set of data. These tools are the combination of algorithms and analytics

A. Big data Analytics

The term 'Big data analytics' itself says the meaning analyzing data which are in a large volume. These large amount data from wide variety of sources such as Social Medias, Sensors, etc. Traditional systems for analyzing which can't handle such large amount of data. And so, specialized software programs are used for these Data Analytics process to handle large amount of unstructured data. Traditional data bases and data ware houses are not sufficient to handle such large volume of unstructured data. Technologies such as Hadoop, Map Reduce, No SQL databases had been emerged to rectify the problems of handling unstructured data sets which are all available in open source too. The purpose of the

978-1-5386-3521-6/18/$31.00 ©2018 IEEE

Survey On The Solution Of Data Skew In Big Data

M. Jenifer
Master of Engineering, Computer science Department,
Sathyabama University,
Chennai, India
refinejm@gmail.com

Dr. B. Bharathi
Professor, Computer science department,
Sathyabama University,
Chennai, India
bharathi.cse@sathyabamauniversity.ac.in

Abstract— **The use of big data is increasing day by day. The problem, while processing big data also increases. Big data processing is done by a powerful tool called Map Reduce. The main problem faced while processing the big data is data skew. In the Past, while utilizing the MapReduce many problems occur like theta join processing, running the application in cloud based environment, execution time being slow due to complex record usage etc. This paper elaborates the problems and solution found by authors while processing the big data using MapReduce tool.**

***Index Terms*—map Reduce, big data, theta join**

I. Introduction

In recent years, many people are coming across the word "Big data". In this internet world, whatever we type in any search engine the data is shown in a fraction of a second. This is due to existence of big data. Big data is processing of large complex data both structured and unstructured. Structured data are data residing in a fixed column with data. Examples are relational database and Excel sheet. This structured data is managed by SQL Queries. Unstructured data is one which does not have fixed column. Most of the data is unstructured like pdf, word document, images, and videos etc. These data are difficult to manage. Some data are semi structured which follows both the rules of structured as well as unstructured. Photos or other graphics in social website can be tagged with keywords such as the creator, date, location and keywords which allow the other user to view, to organize and to locate images mainly used by Face book. XML and other markup languages are being used for semi-structured data.

The main characteristics of big data[13] are volume (size of the data), variety (Type of data being used) and velocity (processing speed of the data). If the size of data increases, the complexity in processing the data and time to complete the execution also increases. The type of data being used and speed of processing the data also counts for the complexity in processing. For example, if the data being used is in plain text with size of above 5 GB and the processor used is able to process only 20 MB at a time then the user has to wait for a long time to process the data and sometimes it may get struck due to huge amount of data.

The data being used is dynamic and vary from time to time. The processing of datasets is mainly done by MapReduce. In a perfect world, all sub-tasks of a parallel task finishes at the same rate within the goal that there is no misuse of resource. But in reality, there will be slight or more variation in execution due to various reasons. Heterogeneous hardware infrastructure, resource contention, and degenerating performance of hardware components are common reasons of such variation in runtime.

The description of below heading is Section 2 describes about the MapReduce framework, section 3 says about the common type of skew effect and section 4 describes what are the common skew occurring in both Map and reducer phase and section 5 briefs about the explanation about few papers dealing with skew and how those authors have handled it.

II. MapReduce Framework

MapReduce [1][12] is a programming model and a related execution for preparing and producing expansive information sets with a parallel, disseminated calculation on a group. Implementation MapReduce framework [14] is done by the following:

- Apache Hadoop
- CouchDB
- Disco Project
- Infinispan
- MongoDB
- Riak

The MapReduce framework is a master – slave concept where the jobs are tracked by the Master. The master is responsible for scheduling the tasks, monitoring them and re-executes the failed tasks for the slaves. The slaves blindly accept the instruction given by the master.

In MapReduce framework, jobs are usually splits in the form of data-set which are further divided into independent chunks and are processed by the map tasks in a parallel manner. The framework sorts the outputs of the maps, which then act as an input to the reduce tasks. Typically both the input and the output of the job are stored in the form of file-system.

The MapReduce framework operates on (key, value) pairs. The framework views the input to the job as a set of (key, value) pairs and produces a different set of (key, value) pairs as the output of the job, possible of different types.

978-1-5090-4556-3/16/$31.00 ©2016 IEEE

Study on the Impact of Distance and Emissivity Measurement for Condition Monitoring of Electronic Circuit Boards

M.S.Sangeetha and N.M.Nandhitha

***Abstract*—In recent years, infrared thermography is the most widely used technique for condition monitoring of electronic equipments. Thermograph are to be analyzed for hotspot extraction and actual hotspot temperature must be determined. however thermograph of low contrast enhancement must be developed through conventional spatial and spectral domain. Contrast enhancement technique work well on normal images, these technique donot provide satisfactory result for thermograph hence in this research work multilevel contrast enhancement and thresholding techniques are used for extracting the hotspot region. Uncertainty in such systems is due to the impact of emissivity and distance on hotspot temperature. In this work, a successful attempt is made to find the constitutive equation for temperature in terms of emissivity and distance. From the research work, it is found that temperature varies nonlinearly with distance and emissivity.**

***Index Terms*—condition monitoring, emissivity, distance and thresholding.**

I. INTRODUCTION

Condition monitoring of electrical and electronic equipments attracts the attention of the researchers. It refers to periodic monitoring of equipments and suitable decision making at the right time. Infrared thermography has gained importance in recent years as it as a non contact, non hazardous and non-invasive technique suitable for the study of temperature profiles. It uses an IR camera that captures thermal patterns which are mapped into thermographs [1]-[3]. Hot spot or cold spot in the thermograph represents anomaly. Hence it is necessary to study the impact of these physical parameters on temperature and to compensate temperature variations before making a decision. In this work, impact of emissivity and distance on temperature variations in thermographs is studied.

M.S. Jadian (2012) proposed a technique to extract the thermal anomalies in electrical equipment by adjusting intensity of image. In this technique, the original images were inverted by subtracting the image from the highest pixel value.

M.S.Sangeetha is with Faculty of Electrical Engg.,Sathyabama University, Chennai, India (sangeetharm28@gmail.com)
N.M. Nandhitha is with Faculty of Electrical and Electronics Engg.,Sathyabama University, Chennai, India (nandhi_n_m@yahoo.co.in)

In order to highlight the Region of Interest (RoI) the inverted image was further subtracted from original image. ROI was then extracted to identity the thermal anomalies in electrical equipment [4]. T.Nunak et al., (2015) proposed a Thermal Imager (TI) which was an effective tool to measure the temperature to detect the failure in electrical equipments. Three different effects in emissivity were carried out using typical experimental setup. By changing the Field Of View (FOV), it was observed that the emissivity remained constant approximately over a range of viewing angle.

Distance between the detector and the object affected the pixel size. When the type of lens was varied, the thermal image resolution changes. Hence it was concluded that, with an appropriate emissivity, FOV and distance, the proposed imager was the best suitable tool to measure the temperature of hotspot [5]. Jeffli et al., (2015) proposed a method for online temperature measurement of rotating machines to detect the defects. Initially, the rotating machine images were captured without any anomalies and stored in the database (reference image). The image of rotating machines was captured periodically. These images were segmented to extract the hotspot. The temperature of the hotspot was calculated and compared with the reference image. The abrupt change in the temperature indicated the defects in the rotating machine [6]. Manjur Ahmed et al., (2015) recursively constructed fuzzy system to indentify the conditions for thermal abnormalities. However in real time applications, temperature variation is also dependant on emissivity and distance which leads to uncertainty in measurement [7].

This paper is organized as follows. Research database and metadata are explained in section II. Section III deals with image enhancement and segmentation techniques. Quantitative characterization is also described in this section. Relationship between emissivity and distance with the temperature correction is described in Section IV. Conclusion and Future work are provided in Section V.

II. RESEARCH DATABASE

In this work anomaly in the PCB of Universal Serial Bus (USB) studied. In order to perform the research work, a research database is created with a total of 24 thermograph by varying the emissivity from 0.1 to 1 and distance from 0.7m to 1 m. Actual temperature is measured using thermometer. Thermographs obtained at a distance of 1m, 0.9m and 0.7m and emissivity of 0.1 and 0.4 are shown in Figures 1and 2 respectively. From the thermographs, it is clearly understood

978-1-5090-3800-8/17/$31.00 ©2017 IEEE

IEEE Advancing Technology for Humanity

2017 International Conference on circuits Power and Computing Technologies [ICCPCT]

Spectral Statistical Analysis of Low Frequency Coefficients from Diagnostic Signals Depicting MHD Disruptions

T.Thaj Mary Delsy[1]
Research Scholar,
Sathyabama University,
Chennai, India.
delsykennedy@gmail.com

N.M.Nandhitha[2]
Professor, Department of ECE,
Sathyabama University,
Chennai, India.
nandhi_n_m@yahoo.co.in

Rakesh L.Tanna[3]
Scientist & Section head,
Aditya Tokamak Operation Group,
IPR, Bhat, Gandhinagar, India.
rtan.ipr@gmail.com

Joydeep Ghosh[4]
Scientist & Project Leader,
Aditya Tokamak, IPR, Bhat,
Gandhinagar, India.
jghosh.ipr@gmail.com

***Abstract*—Aditya Tokamak is a fusion reactor for obtaining nuclear fusion energy from high temperature, ionized gas called plasma. The magnetic field is used to confine plasma in the shape of torus. A disruption is a violent event that terminates magnetically confined plasma. In a disruption, the temperature drops drastically and heat particles are released from confinement on a short timescale and dumped on the vessel wall, causing damage in proportion to the stored energy. The loss of confinement is associated with the production of runaway electrons, which may also produce damage. In order to mitigate disruption, it is necessary for early prediction of disruption. The signals like Halpha, Hard X ray, Plasma current, Mirnov coil signal, Vloop, Soft X-ray should be analyse for the detection of plasma disruption. From the signal if there is any peak in the Halpha, Hard X ray and the non negative value appears in Vloop signals before the decayed of the plasma current indicates the disruption. There are hard and soft disruptions. Since the hard disruption is dangerous, methods like Artificial Neural Network (ANN) using nmultilayer perception and the Fast Fourier Transform (FFT) where already used to find the disruption in terms of frequency component signals. Since Aditya Tokamak signals are non stationary, the above mentioned methods for stationary signal, so those methods are not providing correct and satisfactory results. In this work, different wavelet transforms like Daubechies, Discrete Meyer, Symlets and Biorthogonal were applied on the signal and the corresponding approximation and detailed coefficients were obtained from all the signals in order to obtain the disruption. On an average, of thirty signals are used from different shots for the analysis. Out of the above mentioned wavelet transforms, Discrete meyer and Biorthogonal wavelet are providing the better results than others in terms of the statistical parameters such as mean, skewness and kurtosis. Mean is minimum and the skewness and kurtosis are maximum in the disruption signal, which is confirmed with its time domain analysis. Discrete meyer and the Biorthogonal wavelet transforms provide the spectral information in contrast to frequency domain approaches like FFT. It provides early information about the hard disruption. Wavelet transform is better than FFT and ANN. Since the range of parameters responsible for disruption is not fixed by this method the analysis can be strengthened with Framelet transform.**

Keywords*—*MagnetoHydroDynamics, Plasma Disruption, FFT, Biorthogonal, Discrete Meyer, Skewness, Vloop.

I. INTRODUCTION

In Aditya tokamak, controlled thermonuclear fusion is used for generating electricity by producing and confining plasma within the chamber. Due to unprecedented reasons, disruptions occur in tokamak. Disruptions are of two types namely soft disruption and hard disruption. In hard disruptions, plasma confinement time is reduced drastically and plasma disintegration occurs. Also hard disruptions provide a large amount of energy on the plasma facing components thereby deteriorating the performance of Tokamak machine[1, 2]. It results in both surface and bulk damage namely high erosion losses and large rise in temperature respectively to the materials in Tokamak[3, 4]. In most cases, disruptions are preceded by a large period of MHD signals[5]. Uncontrollable growth of internal MagnetoHydroDynamic instability leads to disruptions in plasma[6]. Hence it is necessary to perform a detailed study on the factors leading to hard disruption. It can be done by acquiring the various signals and analysing them. From a detailed literature survey it is found that Mirnov signals, Hard X-ray signals, Soft X-ray signals, Plasma current, Temperature, H Alpha emission and Loop voltage are the signals used for analysing disruptions in Tokamak[7]. Generalised Single Value Decomposition (GSVD) is used for removing noise from soft x-ray signals to understand the transient response of the signals[8]. There is a sudden, unexpected heavy loss of energy when beta increases beyond the threshold. Single Value Decomposition is used for determining the dominant frequencies from the soft X-ray

978-1- 5090-4967- 7/17/$31.00 © 2017 IEEE

Sensor Node Failure Detection using Check Point Recovery Algorithm

Sandeep Saurav Singh
U.G Student, Department of IT
Sathyabama University
Chennai, India
sandeepsauravsingh@gmail.com

Y. Bevish Jinila
Assistant Professor, Department of IT
Sathyabama University
Chennai, India
bevish.jinila@gmail.com

***Abstract*—Information plays an important role in today's world. For transmitting this information, we use a lot of different kind of network system. Among this the most widely used medium is the wireless sensor network. But due to the high end traffic of the transmission of this information, some failure takes place. Detecting this wireless sensor network failure is very important. For detecting these failure, round trip delay and path algorithm in wireless sensor network has been proposed. But due to the use of large numbers of portable sensor nodes in WSN, probability of sensor node failure gets increased. The fault detection analysis time will increase exponentially with increase of sensor node in WSNs. The objective of executing this project is to resolve these vulnerability present in these sensor nodes. Here we are trying to detect the malicious node so that the data are not transferred to those nodes in order to reduce the data loss, thus enhancing the throughput. The throughput is also enhanced by reducing the number of transmission nodes. By doing this, we can analysis the faulty nodes in a faster manner and can rectify them for better and efficient data transmission**

Keywords*—*Wireless sensor networks(WSNs); malicious nodes; throughput

I. INTRODUCTION

Wireless Sensor Networks (WSN) is now-a-days used everywhere for varied applications. Working with sensors helps in analyzing many fields such as home security, surveillance, industrial monitoring, hospitals, environmental monitoring and military operations and hence reduces human effort. Using large number of sensors in any application increases the Quality of service (QoS).Rapid growth of electronics fabrication has made it possible to manufacture sensors at low cost. On the other hand using large no of sensors also leads to failure hence affecting the QoS. Sensor Node failure can happen due to many reasons such as failure of battery, environmental effects, hardware or software disorder.

For enhancing QoS it is important to detect the failure node and transfer the data associated with that sensor node.Existing system uses Round Trip Delay (RTD) time for selecting the failure node. RTD is the time taken to receive the packet at receiver's end and getting back the acknowledgement to the sender. If the RTD is exceeding its general value then it is expected that the node has failed. This method does not lead to the failure detection at appropriate time. Hence, it can cause data loss in the network.

To overcome this limitation, in this paper the Check Point recovery algorithm (CPRA) is used for fault detection. The CPRA calculates the energy level of each sensor node. At some time interval all the nodes will send a heartbeat which will lead to know that the node is working properly. When the energy level of the node will decrease it will not send any heartbeat to the checkpoint which will let us know that the node is about to fail. So, the data associated with the node can be transferred to some other node which will lead to an increase in the reliability of the system. For the verification of the scalability of this concept, WSNs with large numbers of sensor nodes are implemented and simulated in network simulator NS2.

The rest of the paper is organized as follows. Section 2 briefly discusses about the related works. Section 3 gives a detailed framework of the proposed system. Section 4 discusses about the experimental discussion and evaluation. Section 5 concludes the proposed work.

II. RELATED WORKS

This section addresses the existing related literatures on sensor failure node detection. Ravindra et al. [1] has presented a framework in which the Round Trip Delay (RTD) time of discrete round trip paths are used for the detection of faulty nodes and are compared with threshold value. The limitation of this scheme is that the detection of nodes happen after the expected RTD exceeds. By that time the node already use to be in failed state. Hence, the data associated with that node cannot be retrieved again which leads to packet loss.

Kewei Sha et al. [2] has proposed a state-of-the-art survey of existing multipath routing protocols for WSNs, which are classified into three categories, non-infrastructure based, coding based and infrastructure based, based on the special techniques used in building multiple paths and delivering sensing data.

Peng Jiang [3] has proposed a system in which an improved distributed fault detection (DFD) scheme is proposed by defining new detection criteria. In DFD scheme the failed nodes are checked out by the exchanging data and mutual testing among neighbor nodes in this network. But the

978-1-4673-9802-2/16/$31.00© 2016 IEEE

2017 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC)

Self-Tuning Service Provisioning for Decentralized Cloud Applications

G.Pranai Sai Reddy,
Sathyabama University

Dinesh Kumar Gangipamula,
Sathyabama University

Dr.V. Ulagamuthalvi(Professor)
Sathyabama University

Abstract- Cloud computing has been a massive change in the world of information technology as well as e commerce. The storage facility available with cloud has been a handy tool for all the services provided to the customers. Previously, developers used to buy the resources in a traditional way which led them to high expenses, time complexity. In traditional method of buying and selling a product, the seller fixes a price and buyer has to agree with it when there is scarcity of that particular product and also if a seller can't find an effective buyer whose price matches with him. These cases can lead to dissatisfaction of either buyer or seller. In this project , we came up with an idea by using cloud computing to create a new method of bidding auctions which provides a common platform for both buyers and sellers to trade their storage spaces with their satisfactory prices so that a buyer can find a best seller for his/her budget and similarly a seller can find a best buyer for his/her price. In existing technique, all the methods are developed using Resource allocation algorithm whereas the new proposed technique uses Cloud Resource Allocation algorithm.

INTRODUCTION

Cloud resources are distributed all across the globe. These resources may be owned by multiple users as per their policies and usage. A large scale of buyersand sellers coexist. Based on demand and supply, allocation of resources and claiming them back can happen at any time. Here, we try to employ a faithful dynamic work flow manner which is based on a new technique called cloud resource allocation in which the resources are made much more available to the users. It is majorly used for Online Bidding and Product purchasing based on Auction Mechanism in a faithful dynamic manner. However, we can also implement the online shopping web application containing many products. Here, all the information about providers and consumers will be stored in cloud. Absence of data owner is possible. This new technique is highly expandable and is more efficient in many times and conditions. There will be some cases where there the demand for the product will be high than the supply of it. In this project , we came up with an idea by using cloud computing to create a new method of bidding auctions which provides a common platform for both buyers and sellers to trade their storage spaces

978-1-5090-4324-8/17/$31.00 ©2017 IEEE

Renewable Energy Source based Γ- Z source inverter interlinked to grid with Harmonics optimization Techniques

Catherine Amala Priya. E
Research scholar, Dept of EEE
Sathyabama Institute of Science and Technology,
Chennai, India
rahulcathy@gmail.com

G. T. Sundar Rajan
Associate Professor, Dept of EEE
Sathyabama Institute of Science and Technology,
Chennai, India
Sundarrajan.eee@sathyabama.ac.in

***Abstract*- Γ- Z source inverter entwined between renewable energy source and grid has been proposed in this work. The Key purpose of this work is to improve the voltage at the corresponding boost/ buck levels. To achieve the improvised voltage, harmonics has to be reduced. The reduction in the harmonics in the proposed Γ- Z source inverter has been carried out using two control techniques- Pulse width modulation (PWM) and Space vector pulse width modulation (SVPWM). A comparative study for the optimal control technique has been made between both the techniques. Simultaneously components will be reduced, coupled transformer has been employed which will provide better gain with reduced number of turns and improved modulation ratio. Testing of the proposed model has already been carried out experimentally and the results have proven the reduction of harmonics resulting in better gain.**

Key Words- Γ- Z source inverter, renewable energy source, PWM, SVPWM.

I. INTRODUCTION

With the increase in demand for renewable energy source the need for interface between the domestic consumer and renewable energy source has also increased very rapidly. Inverters being the pioneer in the field of renewable energy conversion, consumers are looking for optimized techniques with improved performance factors.

Initially conversion was a tedious process which involved chopper circuits for achieving stability further a two step process which requires voltage source inverters and current source inverters for obtaining the power, to overcome all this [1] came with the idea of ΓZ source inverters, which proved to have minimized voltage spikes and increase in voltage gain [2] Proposed closed loop model for the ΓZ source inverters with individual closed loop for DC side and AC side, through which accuracy can be increased and system performance can be controlled.

Between a DC sub grid and AC sub grid ΓZ source inverter was used as a inter link in [3] and Droop control strategy was used to obtain an overall control over the DC sub grid and AC sub grid. An Asymmetrical Bi-directional ΓZ source inverter which can perform both buck/ boost operation with the capability of increasing the voltage gain simultaneously with reduced duty ratio [4]. A novel idea was presented in [5], in which zero current transition technique was discussed added transformer usage was avoided instead and 6 switch full bridge topology was used which led to the decrease in leakage current and rapid increase in the gain and efficiency.

Basic boost converter topology with active z type network were proposed and results were acquired in continuous conduction mode and discontinuous conduction mode. Increase in the voltage gain when boosted with this z type networks, which had the input has photovoltaic cells [6].

The proposed ΓZ source inverter overcomes the cost problems when compared with [7], since it uses both interleaved fly back converter and full bridge inverter due to which component usage is increased, conversion process gets tedious and multi stepped added synchronization has to be maintained for the control. [8] & [9] discusses about the chemical properties and physical properties of the solar panel to improve the input efficiency, Cu/Sn–Zn Solder/Ag Interfaces used to reduce the stress between the resistances in solar panel [8] while [9] proposes dynamic voltage restorer otherwise called as six port converter to overcome the sever grid faults caused.

Virtual capacitor along with inverter was used to overcome the voltage ripple and frequency ripple caused by the DC components, while experimental model validated up to 10KVA [10], while the proposed work of ΓZ source inverter has no voltage limitation. Converter-Inverter- Transformer interlinked photo voltaic based grid system medium level voltage utility operated on cascading several such structures, usage of transformers and DC components may increase the junk factors like ripples and produce distortions in the grid, with the use of inductors those have been reduced [11].

Power generation for grid from hybrid renewable energy source with analysis and operation based on converter- inverter topology with harmonic correction which traditionally uses selective harmonic elimination was alternatively replaced by space vector modulation technique and a reduced duty ratio of up to 0.25 using class DE type inverters are included for the references of the proposed work [12-14], since reduced duty ratio with improved gain will lead to better efficiency.

ΓZ source inverter with the duty ratio of 1 to 2, between which by modifying the turn's ratio the gain can be increased, lesser the winding turns higher the gain will

978-1-5386-1508-9/18/$31.00 ©2018 IEEE

2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)

Decentralised Applications Using Ethereum Blockchain

R. Aroul Canessane,
School of Computing, Sathyabama Institute of Science and Technology,
Chennai, Tamil Nadu
aroulcanessane@gmail.com

N Srinivasan,
School of Computing, Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu

Abinash Beuria,
School of Computing, Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
abinashbeuria1@gmail.com

Ashwini Singh,
School of Computing, Sathyabama Institute of Science and Technology,
Chennai, Tamil Nadu
ashwinidotx@gmail.com

B. Muthu Kumar,
Department of Computer Science and Engineering Chennai Institute Of Technology Chennai, Tamil Nadu.
aroulcanessane@gmail.com

***Abstract* — The concept of blockchain without a doubt is a revolutionary concept. It is the underlying Technology behind bitcoin and many more cryptocurrencies. Although the people's focus being only at blockchain as cryptocurrencies in everyday services to do payments online without the interference of a third party will try to replace the current method of cash which is a really slow and ancient method. Blockchain is a zero trust network and this makes it a very powerful tool for various services provided that people are ready to believe and invest in it. In the Ethereum world, the blockchain runs on smart contracts which are self-executing applications that come at a cost of security. This zero-trust network is capable of replacing many of the debated process or activities in our day to day life.One of our biggest concerns is an E-voting system which must be secure. Blockchain being an immutable and append only ledger will not allow for any tampering while also being fully transparent. In this paper, we have implemented and tested a sample e-voting app running as a smart contract for ethereum network using E-Wallets. After an election is held, eventually, the ethereum blockchain will hold the records of ballots and voters thus giving us a clear and trusty network where mishandling is to a minimum..**

Keywords* — *Blockchain, Etherenm, Bitcoin, Smart Contracts, Solidity, e-Voting.

I. INTRODUCTION

Blockchain Technology is becoming more and more evident in our day to day lives. More and more companies have started using blockchain as an underlying network for their daily transactions. When blockchain became popular, it was only used for handling payments using the cryptocurrency Bitcoin, but over the years, various studies have been carried upon and have started to suggest that blockchain can be used in many more areas.

In the ever-growing world, where several transactions need the approval of a third party vendor or can only be done in their presence is becoming a slowing factor in these negotiations. In blockchain, there is no need for a central authority to approve of these transactions or execute the operations. Blockchain is a peer to peer based network which is run by all the nodes that are participating. Because of this, not only we have a zero trust network but also all the structural information is kept within the blockchain network. The devices connected to the network are called nodes and they keep a copy of the blockchain and all of the transaction history in the form of a distributed ledger. To make a transaction on the blockchain, a node must have some incentive for the other nodes to trust this node. This incentive can be in the form of the blockchain's money; Ether gas in the case of Ethereum[2] blockchain. The gas value is required to execute, publish or make transactions on the Ethereum blockchain. Ethereum blockchain consists of smart contracts which reside in the data logic layer of the blockchain. Nodes can read other smart contracts and can read and execute them. Each smart contract needs to have a maximum cost of execution which lets the fellow nodes the know how much gas is required to execute the particular smart contract so that it gets uploaded to the network and be available to every other node.

As mentioned, with the blockchains unique function[8] and a distributed ledger, tampering without setting off an alarm is really tricky. A fork in the blockchain is referred to a change that occurs in the blockchain which causes it to split into two paths. This change can be because of a change in the blockchain's protocol or a forced forgery. Although forks can be accidental or intentional, which can happen because of various reasons, all the nodes agree to a consensus and try to fix the problem in the chain, this is called a hard fork where the participating nodes in the blockchain come to a collective conclusion and agree on making only one of the split nodes as the correct one while the other is discarded and all of its transactions are lost.

978-1-7281-1599-3/19/$31.00 ©2019 IEEE

Cancer Diagnosis System for Classifying Clinically Actionable Genetic Mutation

Y. Ulagamuthalvi[1]
Professor,
School of Computing,
Sathyabama Institute of Science and Technology,
Chennai.
ulagamv@gmail.com

G.Kulanthaivel[2],
[2]Professor,
Department of Electronics,
MITTTR,
Chennai-113.
gkveldr@gmail.com

Idury Jnana Bhaskar Akhil [3],
Jyothindra Pavan Kondepati[3],
[3]Dept of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
Chennai.

Abstract

Machine learning which is sub branch of Artificial Intelligence as helped many medical diagnoses in the form of disease prediction or detection and further diagnosis or prognosis. Generally, for healthcare of cancer diagnosis data scientists use medical records like X-Ray reports and other text-based reports or they use images of different scanning tests. In this system, we are going to use Gene as a primary key and we are going to predict whether there can be a chance of cancer with the mutation or variation on that particular gene. Some variations or mutations can lead to cancer. Lot of research has been done on this by molecular biologists to identify which variation can lead to cancer. So, for a given gene and a variation it is getting hard to predict whether he/she may have cancer by the pathologists. We are helping them by automating the work of referring to the literature of the research done using machine learning.

Key words: ***Gene Variation, Cancer Diagnosis, Data Mining, Machine Learning.***

I. INTRODUCTION

Machine learning is employed in the healthcare industry long back and helping the doctors as well as patients by decreasing the time taken for disease diagnosis or prognosis and prediction or detection. Machine learning is also capable for post diagnosis as well like survival, recurrence. But as data is becoming complicated to understand and to draw the insights lot of work is going on. As so much noise is being present in the data, it is hard to use It in healthcare as accuracy is primary goal in this field.

Generally, Data scientists use X- ray reports and other medical reports in the text form or they will use scanning tests images as data. As, there are many outliers in this data, prediction as well as diagnosis accuracy is decreasing exponentially. Here, every human contains distinct gene. While growing up generally, it may undergo with a variation or mutation. Molecular biologists found that some variations are leading to various cancers. So, lot of research has been done on this and thousands of research papers are being published every year.

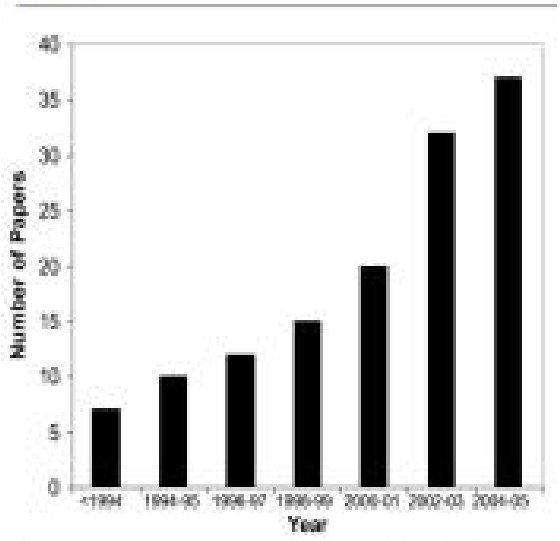


Fig. 1.a Statistics of the publication year 1994 to 2005

978-1-7281-1599-3/19/$31.00 ©2019 IEEE

Authorized licensed use limited to: University of Exeter. Downloaded on May 08,2020 at 03:49:55 UTC from IEEE Xplore. Restrictions apply.

An overview on Change Detection and a Case Study Using Multi-temporal Satellite Imagery

N. Anusha [1,*],
Research Scholar,
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
OMR Road,
Chennai, India
anusha.nallapareddy@gmail.com

B.Bharathi[2]
Professor,
Department of Computer Science and Engineering,
Sathyabama Institute of Science and Technology,
OMR Road,
Chennai, India
bharathivaradhu@gmail.com

Abstract*—Satellite imagery based change detection plays an important role in analyzing the after effects of natural disasters, detecting the changes in city limits due to rapid urbanization, updating the map database, monitoring the factors impacting agriculture, etc., The remote sensors mounted on satellites or aircrafts absorb the light reflected by the earth's surface. The output of these sensors will be a digital image which represents the scene being perceived. In order to extract the useful information from these images, various image processing techniques need to be employed. In this paper, a detailed outline of the steps and various techniques used for detecting the changes in multi temporal remote sensing images is discussed and a case study is done by taking multi-temporal Landsat-8 images covering Hyderabad city. Image differencing method is applied in order to find the changes in the Hyderabad city limits over 2013December and 2017 December time periods.

***Keywords*— Change detection, Remote sensing, Urbanization, Region of Interest (ROI), Filtering**

I. INTRODUCTION

Remote sensing refers to gathering or sensing of information about an object without making any physical contact [1]. Remote sensing has many applications in various fields like military, intelligence, commercial, economic, land surveying, humanitarian, geography, etc., Geo Informatics System (GIS) is the most important remote sensing technique. Primary factors in RS are energy source, wavelength and interaction mechanisms. The data collected by the RADAR during floods will be helpful, as it is not affected by atmospheric interference and cloud cover. Optical remote sensing utilizes visible, short-wave infrared sensors to capture the images of different parts of earth. In remote sensing, the Area of Interest (AOI) is captured with help of a satellite or aircraft or from a certain height from the ground level. It takes considerable time to grasp or analyze the damages/changes with human eye. Therefore the solution is digital image processing.

Image Processing includes image enhancement, image degradation, image segmentation, image classification. This has vast application in the field of medical imaging, military, biometrics, satellite imagery, etc. Image refers to interpretation of images[2]. In remote sensing system temporal resolution is the imaging of the exact same area from the same viewing angle at different periods of time. Sensor's temporal resolution depends on factors like swath overlap, latitude and sensor/satellite capabilities. For example, when a natural disaster like earth quake/flood has occurred, then how much area is affected, how much loss occurred and what measures can be taken to help the people in affected area. Detecting changes in the features of the spectral characteristics only can be done by frequently collecting and comparing the imagery at different intervals of time [3].

Using logistic regression addressed techniques for segmentation and supervised classification of multitemporal multispectral SPOT 5 satellite images. Based on the heights of the buildings, the changes were found. This method fails to detect if any new buildings are constructed with same height or sharp edges, mostly suitable for industrial areas [4]. Conditional Random Fields (CRF) model which is a supervised method is used for contextual classification. The dataset/images taken for experimenting are limited to medium resolution [5]. Using unsupervised classification proposed a framework for detecting changes in buildings by utilizing the height and spectral features information. The method applied is a complicated method. This method failed to detect few buildings and changes present in those undetected buildings [6].An automatic change detection method is proposed to detect changes in high resolution multispectral images collected by ZY-3.Two threshold groups are selected instead of using single threshold point in order to enhance the quality of training samples [7]. Hierarchical Change Detection Tests (HCDTs) which is one of the effective online algorithms is presented. It is composed of two hierarchical layers, detection layer and validation layer, which help for detection of changes in data streams. While segmenting the images, few buildings are lost in the process [8].A novel change detection method called super pixel-based change detection method is addressed for extracting features from the very high resolution multi-spectral image[9]. The Table I below gives the Comparison of previous studies.

DEPARTMENT OF ARCHITECTURE
NATIONAL INSTITUTE OF TECHNOLOGY
TIRUCHIRAPPALLI | TAMIL NADU | INDIA

NCACESD 2019

Proceedings of

1st National Conference

ON

Recent Trends in Architecture & Civil Engineering towards Energy efficient and Sustainable Development

10th – 12th January 2019

Edited By
Dr. K. Thirumaran
Dr. L. Saikala

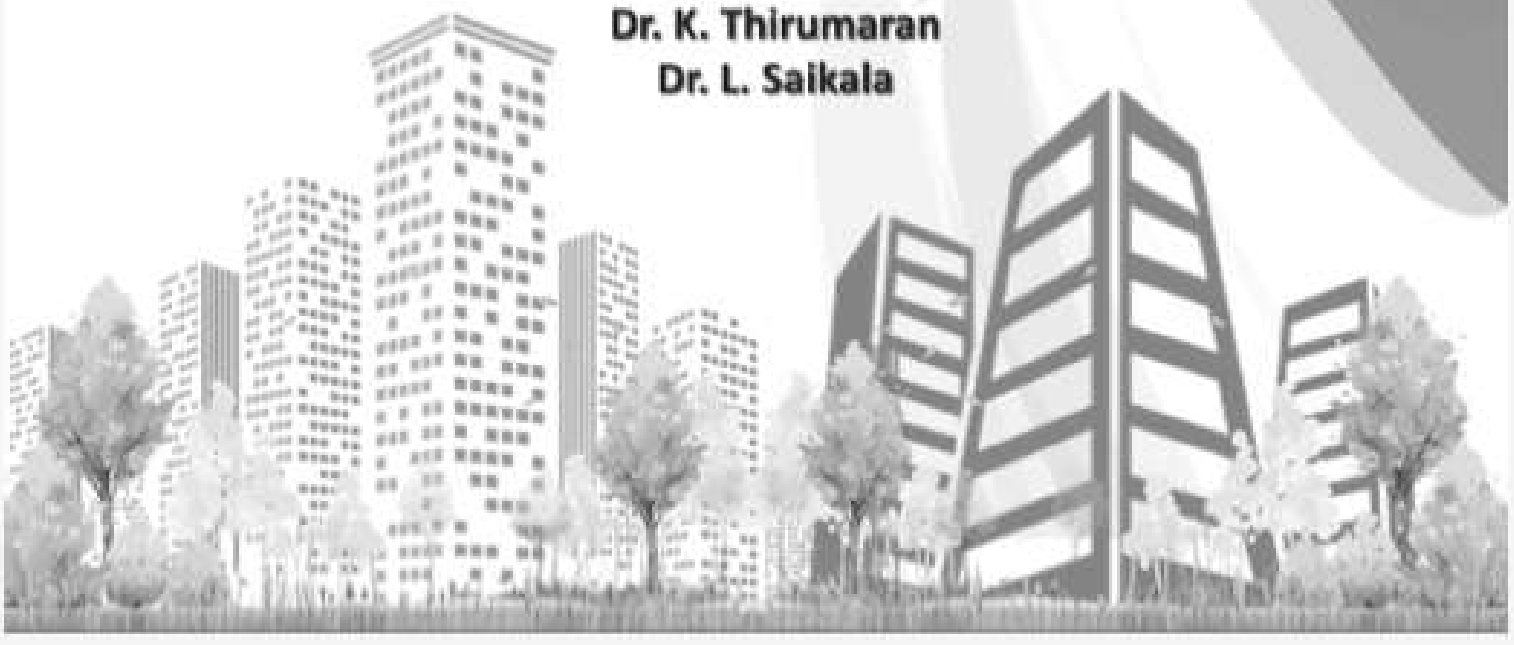


TABLE OF CONTENTS

1	Adaptive Responses to Climate Change - Impact on Water at low-income households in Coastal Tamilnadu *R Rajeswari, Cibi C B & Ranee Vedamuthu*	1
2	Exploring energy efficiency in a guest house complex by implementing ECBC building envelope measures in warm-humid climate of Tiruchirappalli *Niranjana P & K Thirumaran*	9
3	'Express'ing the cultural experience in historic urban centres: People and qualities *Godwin Emmanuel Jeyaraj & Meenakchi Sundaram*	17
4	An investigation on the thermal behaviour of building envelopes using Infrared thermography *S Subhashini & K Thirumaran*	25
5	Urban Morphological Role of Housing Typologies on Energy Demand: A Case of Navi Mumbai *C Ravichandran & P GoapalaKrishnan*	30
6	Understanding the sustainable construction techniques practiced in Auroville and how well they perform by taking the case of a building *Mithra R & Swetha J*	39
7	Mitigation of Adverse Impacts of Flooding in Environmentally Sensitive Areas with Specific reference to Coorg District in Karnataka *Sindhu S &. Vaseem Anjum Sheriff*	49
8	Energy Efficiency & Sustainable Design Strategies in Nahargarh Fort, Jaipur *Manish, Praveen Jangid & Kush Kumar Saxena*	56
9	An investigation of passive strategies adapted in a Government hospital, Chennai *S Swetha, L. Kadambari & S. Kavitha*	65
10	Analyzing the Deciphering Planning Strategies used in Ancient Dravidian Temple Architecture *A Purushothaman & K Thirumaran & A Sivaraman*	72
11	Thermal Comfort Analysis of Rammed Earth in the Indian Context *Prajna Aigal*	82
12	Traditional Sustainable Design Strategies - A Case of Pol Houses of Ahmadabad *Shanu Raina, Bhagyalaxmi & S Madapur*	92
13	"Sense of place" and identity elements alongside the old and the new highways in Tamil Nadu, India *G Ophylia Vinodhini & A Meenakchi Sundaram*	99
14	Assessment of daylight performance in classrooms: To achieve optimal daylighting design strategies for classroom through simulations *Mohammed Shibin N*	109
15	An Assessment Review of Building Professionals towards the Sustainable housing Developments in Nagapattinam Region *A Purushothaman & K Thirumaran*	118

An investigation of passive strategies adapted in a Government hospital, Chennai

Ar. S. Swetha[1], Ar. L. Kadambari[2] & Ar. S. Kavitha[3]

[1,2]*PG Student Sathyabama Institute of Science & Technology, Chennai, India.*

[3]*Assistant Professor, Sathyabama Institute of Science & Technology, Chennai, India.*

Corresponding Author Email ID: shiv.bala99@gmail.com

Abstract: A sustainable, energy-efficient and healthy environment is achieved through passive strategies. In older days, the use of traditional, context specific techniques were developed to improve the indoor thermal comfort. With an intension to explore these approaches, through spatial planning, architectural features, building components, building materials that contribute to improve the thermal comfort conditions, a Government hospital constructed in the early 20th century is identified as a live case study. This paper further investigates how the impact of these techniques has evolved in today's context.

Keywords: passive strategies, traditional approaches, thermal comfort, hospital.

INTRODUCTION

A key to sustainable building is passive design. Passive strategies respond to local climate and site conditions, maximizing user's health and comfort while minimizing the energy consumption. Climate responsive building is designed to adopt itself to the precise area and weather conditions where it is constructed. The movement and location of sun, seasonal changes, and natural features in the surrounding topography, environmental aspects, and climatic records are considered to achieve thermal comfort and wellbeing of the inhabitants.

Chennai located on the thermal equator, experiences hot and humid climate most of the year. May to June has a higher temperature of around 35 to 40 °C, December to January being the coolest months with temperatures around 19–25 °C. Being hot and humid, amount of moisture in the air is high almost throughout the year which causes thermal discount eventually affecting the efficiency of the inhabitants. The average rainfall is about 1400mm; Chennai gets most of its rains from northeast monsoons and cyclones from mid-September to mid-December. The wind direction varies according to the site location and topography.

The ancient traditional construction was an influence of social, cultural, economic and climatic factors and geographical constrains of the region. The vernacular architecture principles of Chennai were derived from various observations and experimental knowledge gathered over the past several centuries.

The highly efficient strategies for providing thermal comfort in all seasons were based on natural and passive techniques in Traditional Colonial architecture. Influence of British rule in Chennai was evident in the architectural development during that period to facilitate the need for administration and health care. This paper is an attempt to study the old traditional approaches and passive strategies in architectural heritage building of early 20th century. The old Regional Institute of Ophthalmology and government ophthalmic hospital Egmore.

METHODOLOGY

This paper illustrates the vernacular architecture of hospital building. The study comprises of 2 phases, quantitative and qualitative analysis emphasizing on thermal comfort.

Quantitative analysis

On site measurements and readings were taken to study the thermal comfort of the Lady Lawley block. Air temperature and relative humidity (Thermo hydrometer), wind movement (Anemometer), natural illuminance (Lux meter) were measured for 12 hours, in 8 fixed locations at an interval of one hour, for a period of 2 days. The instruments were kept at 1.5m height from ground level and 1m away from the walls and surfaces. Weather data was collected from Nungambakkam meteorological station as a reference for the study.

AIP Conference Proceedings

Mechanical and corrosion behaviour of $Gd_2Zr_2O_7$ – HAp and $Sm_2Zr_2O_7$ – HAp composite coated Ti-6Al-4V alloy for biomedical applications

Cite as: AIP Conference Proceedings 2311, 080001 (2020); https://doi.org/10.1063/5.0034655
Published Online: 07 December 2020

Gobi Saravanan Kallaraj, Anderson Arul Gnana Dhas, Karthik Alagarsamy, A. M. Kamalan Kirubaharan, Shaik Sarwath Hussain, and K. Sasitharan

ARTICLES YOU MAY BE INTERESTED IN

Computational analysis on hybrid composite material
AIP Conference Proceedings 2311, 080007 (2020); https://doi.org/10.1063/5.0034009

A review on tribological behavior of polymer composite impregnated with carbon fillers
AIP Conference Proceedings 2311, 070030 (2020); https://doi.org/10.1063/5.0035408

Study of aluminium metal matrix composite material characteristics for potential use in disc brake rotor
AIP Conference Proceedings 2311, 080006 (2020); https://doi.org/10.1063/5.0034540

Challenge us.
What are your needs for periodic signal detection?
Watch
Zurich Instruments

© 2020 Author(s).

Mechanical and Corrosion Behaviour of $Gd_2Zr_2O_7$ – HAp and $Sm_2Zr_2O_7$ – HAp Composite Coated Ti-6Al-4V Alloy for Biomedical Applications

Gobi Saravanan Kaliaraj[1], Anderson Arul Gnana Dhas[2,a)], Karthik Alagarsamy[1], A M Kamalan Kirubaharan[1], Shaik Sarwath Hussain[2], K. Sasitharan[2]

[1]*Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, India*
[2]*School of Mechanical Engineering, Sathyabama Institute of Science and Technology, Chennai, India*

[a)] Corresponding Author Email: chrisson100@gmail.com

Abstract. Cytotoxicity is a crucial problem in biomedical implant applications which could be overcome by enhancing the corrosion resistance through surface coating technique.Studies on biological properties and bio corrosion and a supported study on microstructureof plasma sprayed $Gd_2Zr_2O_7$ hydroxyapatite and $Sm_2Zr_2O_7$ hydroxyapatite coatings on type titanium (Ti-6Al-4V) of grade 5 samples are carried out in this research work. The samples are tested for Bio corrosion behaviour in Artificial Blood Plasma (ABP) as medium to study the corrosion behaviour of ceramic coated titanium (Ti-6Al-4V). The results indicated that the thick and non-porous oxide passive layers formed over the ceramic coated samples augmented the corrosion resistance than bare titanium (Ti-6Al-4V) substrates. The tribology test was performed for the coated samples.

INTRODUCTION

Biomaterials invariably get in touch with the living tissues of organisms[1,2]. The compatibility has to be studied when introducing the biomaterials in parts of human being[3-5]. In the present scenario implant material such as 316LSS, Titanium and its alloy are playing major role in implant industry. Among this Ti-6Al-4V alloy is having good corrosion resistance, low density, biocompatibility and mechanical strength. Due to this nature of Ti alloy we have selected this material for our study. Even though it also having some disadvantage like poor wear resistant and less antibacterial property we need to go for some surface modification. To enhance these properties, we have introduced bio ceramic coating with HAp [HAp, $Ca_{10}(PO_4)_6(OH)_2$], $Gd_2Zr_2O_7$ and $Sm_2Zr_2O_7$.Gadolinium (Gd) and samarium (Sm) for our purpose of study. Further we have mixed both the Gd and Sm with Zr to improve their biological properties. Recently ZrO_2 has occupied its place among bio ceramics in implant industries because of its high mechanical strength, fracture toughness and good cytocompatibility with living tissue[6,7]. On the other hand, $Gd_2Zr_2O_7$ This material has good performances in mechanical properties (strength and hardness), thermodynamic stability and leaching durability, and also has high resistance to radiation damage[8,9]. The modification HAp-$Gd_2Zr_2O_7$ / $Sm_2Zr_2O_7$ composite coatings were achieved by plasma spray coatings. In the present study we focused on plasma sprayed HAp- $Gd_2Zr_2O_7$ / $Sm_2Zr_2O_7$ composite coating. The coated samples were characterized through XRD to confirm the phase of the coated materials. Corrosion and tribological behaviour of the coated samples were analysed through electrochemical impedance spectroscopy (EIS) tribometer.

3rd International Conference on Frontiers in Automobile and Mechanical Engineering (FAME 2020)
AIP Conf. Proc. 2311, 080001-1–080001-5; https://doi.org/10.1063/5.0034653
Published by AIP Publishing. 978-0-7354-4049-4/$30.00

AIP Conference Proceedings

Microstructure and corrosion behaviour of Ni-Ag and Ni-Cu composite coating for industrial applications

Cite as: AIP Conference Proceedings 2311, 080009 (2020); https://doi.org/10.1063/5.0034166
Published Online: 07 December 2020

G. Senthilkumar, S. Bharadwaj, R. Akhil, Gobi Saravanan Kaliaraj, Nivin Joy, and S. Ganesan



© 2020 Author(s).

Microstructure and Corrosion Behaviour of Ni-Ag and Ni-Cu Composite Coating for Industrial Applications

G. Senthilkumar[1,a],S. Bharadwaj[1,b], R. Akhil[1,c], Gobi Saravanan Kaliaraj[2,d], Nivin Joy[1,e], S.Ganesan[1,f]

[1]*Departmnet of Mechanical Engineering, Sathyabama Institute of Science and Technology, Chennai – 600 119, India.*

[2]*Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai – 600 119, India.*

[a)] Corresponding author: tosenthilgs79@gmail.com

Abstract. Corrosion is the characteristic degradation of materials by interaction with the operating environment. One of the important and widely adopted method to prevent corrosion is coatings on the materials. The intention of the present study is to understand the corrosion behaviour of Ni-Ag and Ni-Cu coatings using DC/RF magnetron co-sputtering technique onto stainless steel (Grade 304) substrates. . Thin film X-ray Diffraction method to decide the crystalline structure and provide information on unit cell dimensions. Field Emission Scanning Electron Microscopic method (FESEM) to produce a magnified image with high resolution. Atomic Force Microscopic method (AFM) to study the surface properties of the coated material. Electrochemical corrosion test to analyse the corrosion pattern of the coated materials. The results suggest that the surface modified SS304 exhibits better corrosion protection as compared with uncoated samples.

Keywords: Corrosion, Coatings, DC/RF Magnetron sputtering, Ni-Cu alloy,Ni-Ag alloy, Industrial applications

INTRODUCTION

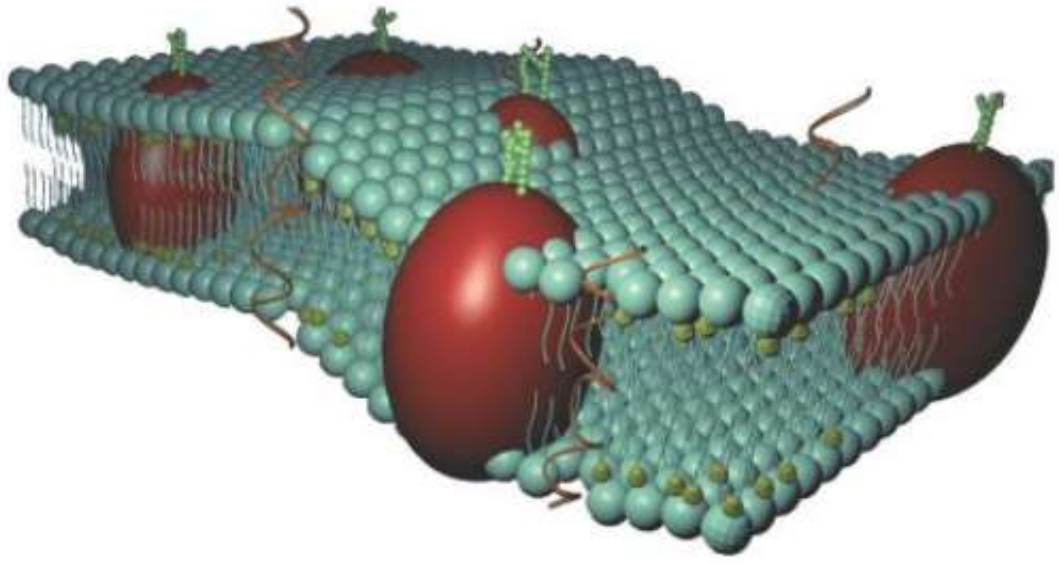
Stainless steel is used in industries where working temperatures are high. Special grade steels have wider applications in sewage plants, desalination plants, oil and gas industries. Easy-clean and hygienic quality of stainless-steel gives scope for medical applications as well. In order to improve the resistance property of stainless steel at high temperatures and marine environment, Ni-Cu and Ni-Ag composite coatings are applied on stainless steel. Ni-Cu alloys have wide industrial and marine applications. Cupronickels are best known for its corrosion resistive property and impressive thermal conductivity. Pipelines for seawater, heat exchangers, powerplants, condenser material and tubing in ships are some of its common applications [1-3]. In addition, Ni-Ag is used as a biofilm inhibition coating and antimicrobial coating over stainless steel. Cupronickel alloys have failed to last long in marine environments and most failures occur on the surface. These failures might affect the engineering systems that work offshore and also the working of ships. Improvement of these alloys is necessary in order to increase the corrosion resistance quality. A thin layer of Cu-Ni alloy act as shielding layer which evolved corrosion resistance property due to strong inner barrier Cu_2O layer [4-6]. Further, inner layer of cuprous oxide containing Ni^{2+} and Ni^{3+} and outer layer of Cupric oxide with water molecules were found on the surface of the resistive Ni-Cu alloy coating in sea water environment [7]. Nickel present in the Cu-Ni composite coating improves the corrosion resistance and mechanical strength of the alloy. The impact of Nickel significantly improves by the presence of iron in Cu-Ni alloy. Similarly, silver has many significant applications because of its antifungal, antifouling and antibacterial properties. Silver is a very good conductor of heat and has high electrical conductivity. In most biomedical applications, bacterial attack is a predominant problem that needs to be avoided in its initial stage. However, in marine plants, seaweeds and sponges have the capacity to degrade silver to silver nanoparticles [8]. To enhance the properties of silver, Nickel has been alloyed along with silver over stainless steel. Coating techniques and surface enhancement

3rd International Conference on Frontiers in Automobile and Mechanical Engineering (FAME 2020)
AIP Conf. Proc. 2311, 080009-1–080009-6; https://doi.org/10.1063/5.0034166
Published by AIP Publishing. 978-0-7354-4049-4/$30.00

Lipid Bilayers

Properties, Behavior and Interactions

Nova Medicine & Health



Mohammad Ashrafuzzaman
Editor

NOVA

Table of Contents

Table of Contents

Preface

Chapter 1. Fundamentals of Plasma, Nuclear and Mitochondrial Lipid Membranes
(Mohammad Ashrafuzzaman, Department of Biochemistry, College of Science, King Saud University, Riyadh, Saudi Arabia)

Chapter 2. Mechanical Stresses in the Lipid Bilayer of Erythrocyte Membranes
(Pavel V. Mokrushnikov, PhD, Department of Physics, Novosibirsk State University of Architecture and Civil Engineering, Novosibirsk, Russia)

Chapter 3. The Role of the Lipid Bilayer in the Erythrocyte Membrane Structural Changes
(Ivana Pajic-Lijakovic and Milan Milivojevic, Faculty of Technology and Metallurgy, Belgrade University, Belgrade, Serbia)

Chapter 4. Microbial-Derived Bioactive Lipopeptides: Pore-Forming Metabolites in Lipid Bilayers
(Yulissa Ochoa, Mariana Bernal, and Daniel Balleza, Department of Chemistry ICET, Universidad Autónoma de Guadalajara, Lomas del Valle, Zapopan, Jalisco, México)

Chapter 5. Colchicine Induced Ion Channel Formation into Membranes as a Mechanism behind Chemotherapy Drug Cytotoxicity of Cancer Cells
(A. A. Alqami, S. Zargar and Md. Ashrafuzzaman, Department of Biochemistry, College of Science, King Saud University, Riyadh, Saudi Arabia)

Chapter 6. Molecular Stability Analysis of Reverse Micelles Role in Breast Cancer Drug Delivery and It's Dynamics and Simulation Studies
(Dhivya Shanmugarajan, N. Premjanu, Ganesh Munuswamy, Lakshmi Jayasri Akkiraju, Sushil Kumar Middha and Sureshkumar Chinanga, Department of Biotechnology and Biochemistry, Science Research Centre, Maharani Lakshmi Ammanni College for Women, Bengaluru, Karnataka, India, and others)

Chapter

Carbon Nanotubes: Synthesis, Properties and Applications

Aravind Kumar Jagadeesan, Krithiga Thangavelu and Venkatesan Dhananjeyan

Abstract

Recent discoveries of salient carbon nanoforms have paved tremendous interest among research and also toward their discrete applications in scientific fields. Various generation methods for carbon nanotubes (CNTs) involve chemical deposition of vapor, discharge using electric arc and laser ablation mechanism which were driven by functionalization, chemical addition, doping, and filing such that in-depth characterization and manipulation of CNTs were possible. The in-built elasticity, electromechanical, chemical, and optical properties of CNTs have a notable impact on its stability and reactivity. Perhaps, the flexibility along with its determined strength makes them to validate its potential application in diverse fields which enables that these CNTs will definitely procure a prominent role in nanotechnology.

Keywords: nanotechnology, nanotubes, synthesis, electromechanical, nanosensor

1. Introduction

Carbon materials can be grouped into three classifications based on their period of advancement: classical carbons, nano ones and new carbons. Cracking carbons incorporate engineered graphite squares principally utilized as anodes, carbon blacks, what are more, enacted carbons, for which creation systems were created before [1] the 1960s. During the 1960s, carbon materials not quite the same as these great carbons were designed: carbon filaments from different forerunners, including fume developed carbon filaments; pyrolytic carbons delivered by means of concoction fume testimony forms; glasslike carbons with high hardness and gas impermeability; high-thickness isotropic carbons created by isostatic squeezing; intercalation mixes with various functionalities, for example, high electrical conductivity; and precious stone like carbons as straightforward carbon sheets. These recently evolved carbon materials are grouped as new carbons [2]. Since the 1990s, different fullerenes with shut shell structure, carbon nanotubes with nanometer distances across, and graphene pieces of just a couple molecules' thickness has stood out from nanotechnology; these are ordered as nanocarbons.

On the off chance that these carbon materials are considered from the perspective of their surface, be that as it may, they might be ordered into two gatherings: nano-textured and nano-sized carbons [3]. Most carbon materials in the new carbon classification are delegated nano-textured carbon, in light of the fact that their nano texture is controlled by means of various procedures in their creation, notwithstanding the basic control. Then again, fullerenes, carbon nanotubes, and graphene can be delegated nano-sized carbon, the shell size of fullerenes, breadth

IntechOpen

12. Assimilating SEVIRI Satellite Observation into the Name-III Dispersion Model to Improve Volcanic Ash Forecast
Prajakta Patil, I. M. Watson, Shona Mackie, Prashant K. Srivastava, Tanvir Islam, and Sourabh Sakhare 151

13. Geoinformation Technology for Drought Assessment
Arnab Kundu, D. M. Denis, N. R. Patel, R. K. Mall, and Dipanwita Dutta 171

14. Introduction to Landslides
H. K. Pandey 181

15. Probabilistic Landslide Hazard Assessment using Statistical Information Value (SIV) and GIS Techniques: A Case Study of Himachal Pradesh, India
Ankit Sharma, Ujjwal Sur, Prafull Singh, Praveen Kumar Rai, and Prashant K. Srivastava 197

16. One-Dimensional Hydrodynamic Modeling of the River Tapi: The 2006 Flood, Surat, India
Dhruvesh P. Patel, Prashant K. Srivastava, Sudhir Kumar Singh, Cristina Prieto, and Dawei Han 209

Section IV: Ocean Hazards and Disasters

17. Tropical Cyclone–Induced Storm Surges and Wind Waves in the Bay of Bengal
Prasad K. Bhaskaran, A. D. Rao, and Tad Murty 239

18. Space-Based Measurement of Rainfall Over India and Nearby Oceans Using Remote Sensing Application
Anoop Kumar Mishra and Kishan Singh Rawat 295

19. Modeling Tsunami Attenuation and Impacts on Coastal Communities
S. Piché, I. Nistor, and T. Murty 309

Index 325

18

Space-Based Measurement of Rainfall Over India and Nearby Oceans Using Remote Sensing Application

Anoop Kumar Mishra and Kishan Singh Rawat

ABSTRACT

Water is one of the most common substances in the Earth–atmosphere system. The source of all water in its most desirable state is rainfall. It is very crucial in determining the nature of our globe at a fundamental level and is essential for sustaining life by providing pure water over the land. The economy of India is dependent on agriculture, which in turn depends on rainfall. This chapter describes the various ways of monitoring rainfall. It describes the limitations of the conventional tools to measure rainfall. Spatial and temporal variability of rainfall over India is also discussed. The chapter gives an introduction to the principle of the satellite remote sensing to monitor rainfall. It discusses the limitations of available global rainfall products for their applicability over India. It also describes techniques to monitor rainfall over the Indian region and their validation with available ground truths.

18.1. INTRODUCTION

Rainfall is one of the atmospheric parameters that has a direct effect on our life. While rain brings freshwater to support life on Earth, extreme rain events cause destruction and misery due to flash floods and droughts. The accurate estimation of rainfall at various temporal and spatial scales has many applications in meteorology, hydrology, and climate studies. Rainfall is a very difficult parameter to monitor due to its high spatial and temporal variability. Generally, rainfall estimation has been a clear-cut matter of setting out rain gauges. Conventional measurement of rainfall from a rain gauge is very accurate, but it gives point measurement. Moreover, rainfall measurement by rain gauges is greatly affected by various factors of topography, site, wind, and gauge design. Weiss and Wilson (1958) reported that rain gauge catch is subjective to the presence of nearby objects and structures that may act as shields and hence decrease the catch from what would have been measured in their absence. Wind is also one of the factors contributing to errors in rain gauge observations. Monthly rainfall can be underestimated by about 12% by average winds as light as 5 m/s over a region where half the monthly rainfall comes at rates less than 1.8 mm/h. A very dense network of rain gauges has to be set up to measure rainfall accurately over an entire area, which is very expensive and is nearly unfeasible. Mishra (2013) studied a 50 km × 50 km region over the southern part of India and found that three or less rain gauges over the region are not sufficient to monitor daily accumulated rainfall accurately. The density of rain gauge networks is not adequate to monitor rainfall over the Indian region.

The advent of weather radar offers an opportunity to monitor rainfall. However, in the case of weather radar too, for the measurements of rainfall, there are various limitations that restrict the use of such a system. These limitations include the lack of a proper relationship of backscattered microwave energy to drop size distribution, reflectivity versus rainfall relationship, beam filling problems, attenuation of the radar beam by intervening drops, absorption and reflection by the ground (anomalous propagation), and signal calibration. Moreover, due to its

Center for Remote Sensing and Geoinformatics, Sathyabama University, Chennai, Tamil Nadu, India

Techniques for Disaster Risk Management and Mitigation, First Edition. Edited by Prashant K. Srivastava, Sudhir Kumar Singh, U. C. Mohanty, and Tad Murty.
© 2020 John Wiley & Sons, Inc. Published 2020 by John Wiley & Sons, Inc.

Dr. Rekha Chakravarthi
Dr. N.M. Nandhitha
Dr. S. Emalda Roslin
Ms. M.S. Sangeetha
Dr. A. Jose Albin

Soft Computing Based Hybrid Mining Technique for Effective Retrieval of Audio Signals

Smart Mirror- Digital Magazine for University Implemented Using Raspberry Pi

R.Akshaya,
Department of IT,
Sathyabama Institute of Science and Technology.
Akshaya21997@gmail.com

N.Niroshma Raj,
Department of IT,
Sathyabama Institute of Science and Technology.
Niroshmaraj.n@gmail.com

S.Gowri,
Department of IT,
Sathyabama Institute of Science and Technology.
Gowriamritha2003@gmail.com

***Abstract*—The design of smart mirror using Raspberry Pi. This may look like the regular mirror but the difference is it displays weather, temperature, time, and web application for college. The smart mirror is implemented using peripherals like raspberry pi, LED monitor covered with a two-way acrylic sheet. The microcontroller can connect to the Internet and can retrieve data from the internet which displays weather report of the particular location. This describes how smart mirror is built using the raspberry pi as hardware.**

Keywords—Raspberry Pi, Internet of things, weather, Web Application, Mirror.

I. INTRODUCTION

Everything in this World is becoming Spruce. Why not Mirror? The smart mirror is same as smart television, smartphone etc. The smart mirror is an improvement attempt to reinforce the mirror with magnified features along with weather forecasting, and time similar to the place. Along with displaying the weather report and time we have created a web application for our college. Time is set by the computer, the climate of the particular place is set by API which dynamically changes accordingly. The application is built in such a way that it has two buttons where [1] one is for map and Architecture of our college. By mentioning the start and destination person can able to identify the path inside the campus. [2] Other is the page designed in such a way that it is used for students inside the campus to view mirror as well as get the important details like course registration, Placement and college blogs. The database is used for storing details of course registration that can be accessed only by admin. The smart mirror is built using theraspberry pi and LED Monitor screen covered with Acrylic sheet (two-way reflective sheet).

The Raspberry Pi is credit card sized single-board computers, which plugs into the computer monitor and uses a standard keyboard and mouse. It is a capable little device that enables people to explore computing. [3] The flow of this platform is similar to MVC. Model-View-Controller is a software design pattern for developing a web application. Here the Controller, Admin receives all requests for the course and then works with the Model to prepare any data needed by the View. The View then makes use of the statistics organized with the aid of the Controller to generate a final presentable response to the students.

A. The Model

The model is architecture for saving the statistics of the operation. Its return to the request from college students and it also return to instructions from the admin to amend itself.

B. The View

It approach the presentation of statistics in a selected layout, here it's far a replicate that presentations the records.

C. The Controller

The controller is for executing to the user input and perform interactions on the data model objects. The controller inherits the input, it validates the input and then performs the business functions that modify the state of the data model.

The rest of the paper is prepared as below. Section II presents Literature Survey Section III presents some related work. Section IV Proposed work of the Smart Mirror. Section V Result of the system. Section VI conclusion and future work. Section VII References.

978-1-5386-5743-6/18/$31.00©2018 IEEE

Advances in Intelligent Systems and Computing 940

Ajith Abraham
Aswani Kumar Cherukuri
Patricia Melin
Niketa Gandhi *Editors*

Intelligent Systems Design and Applications

18th International Conference on Intelligent Systems Design and Applications (ISDA 2018) held in Vellore, India, December 6–8, 2018, Volume 1

Reduced Complexity Affine Projection Algorithm Based on Variable Projection Order and Multiple Sub Filter Approach

S. Radhika[1(✉)] and A. Chandrasekar[2]

[1] Sathyabama Institute of Science and Technology, Chennai, India
radhikachandru79@gmail.com
[2] St. Joseph's College of Engineering, Chennai, India

Abstract. The main aim of the paper is to reduce the complexity of affine projection algorithm (APA) by using multiple sub filter together with variable step size and projection order so that there is no compromise in the performance measure. Here the entire filter length is segmented into smaller segments (sub filter) in order to reduce the projection order which is the main cause for the complexity. However the steady state error increases with the use of the multiple sub filter approach. To maintain an overall lower steady state error, a variable step size together with a variable projection order is used. A simple rule for the automatic change of projection order based on the instantaneous value of output error is also proposed. Therefore faster convergence, lesser steady state error criteria is met with the reduction in computational complexity and computational time. Simulation is performed in the context of echo cancellation application to prove the validity of the algorithm.

Keywords: Multiple sub filter · Affine projection · Steady state error · Convergence · Complexity · Variable projection order · Variable step size

1 Introduction

An APA is one of the popular algorithm of the family of adaptive filters which is based on the affine sub space [1] where the adaptive rule is based on the delayed version of the input vector called as projection order. The flexibility in convergence, steady state error and the complexity makes this algorithm a suitable candidate for a variety of applications especially when the input is highly correlated such as speech signal [2]. The length of the input vector, the projection order and the step size were the deciding features for the performance of the algorithm [3]. The steady state mean square error (MSE) analysis of APA indicated that there is a tradeoff between convergence rate, steady state error and computational complexity [4]. Variable step size approach has been proposed for improved convergence rate, lesser steady state MSE [5–7]. The major drawback of this approach is that the complexity reduction is not achieved as the projection order is maintained as a constant value. The final steady state error obtained is more than the optimal minimum as the steady state MSE is directly proportional to the projection order. Multiple sub filter approach for the adaptive filter is proposed [8–10]. This approach is

© Springer Nature Switzerland AG 2020
A. Abraham et al. (Eds.): ISDA 2018, AISC 940, pp. 610–619, 2020.
https://doi.org/10.1007/978-3-030-16657-1_57

Lecture Notes in Electrical Engineering 521

Ganapati Panda
Suresh Chandra Satapathy
Birendra Biswal · Ramesh Bansal
Editors

Microelectronics, Electromagnetics and Telecommunications

Proceedings of the Fourth ICMEET 2018

Springer

Prospect of Pico-Hydro Electric Power Generation Scheme by Using Consuming Water Distributed to Multi-storage Building

P. Velazhagan, A. Ramesh Babu, M. D. Vijayakumar and G. T. Sundar Rajan

Abstract This paper defines the implementation of domestic-based pico-hydroelectric power generation system in the multi-storage building to generate electricity which can definitely minimize the electricity demand of the building. The proposed system implements mini-generator in the path of the main pipeline from the overhead water tank. The kinetic energy present in water flow from the overhead tank is converted into electrical energy. The generated electrical energy is stored in batteries by converting the AC power into DC power by using power electronic converter. It is one of the alternative sources of energy conservation. The system was implemented in a residential building by having AC generator of 5 W. The key aim of this paper is to demonstrate practically that power can be generated from the overhead water tank and can be stored and utilized for the building.

Keywords Pico-hydro turbine generator (PHTG) system
Permanent magnet AC generator · Energy storage · Turbine · Green energy
Domestic consuming water · Renewable energy

1 Introduction

India is a developing country in the world with creating successful outstanding achievement in space, science and technology, nuclear missile technology, civil engineering, and information technology in all discipline. Today India's estimated population is over 1250 million and currently rated a second highest populated country in

P. Velazhagan · A. Ramesh Babu (✉) · M. D. Vijayakumar · G. T. Sundar Rajan
Sathyabama Institute of Science and Technology, Chennai 600119, Tamil Nadu, India
e-mail: rameshbabu.eee@sathabama.ac.in

P. Velazhagan
e-mail: velspriya@yahoo.co.in

M. D. Vijayakumar
e-mail: mdvijay79@yahoo.co.in

G. T. Sundar Rajan
e-mail: sundarrajan.eee@sathyabama.ac.in

© Springer Nature Singapore Pte Ltd. 2019

G. Panda et al. (eds.), *Microelectronics, Electromagnetics and Telecommunications*, Lecture Notes in Electrical Engineering 521,
https://doi.org/10.1007/978-981-13-1906-8_23

S. M. Kumar and L. Lakshmanan, "A Situation Emergency Building Navigation Disaster System Using Wireless Sensor Networks," *2018 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2018, pp. 0378-0382, doi: 10.1109/ICCSP.2018.8524370.

Publisher: IEEE

Cite This

PDF

R. Sathiyavathi ; B. Bharathi All Authors

4
Paper
Citations

280
Full
Text Views

Abstract

Document Sections

I. Introduction

II. Related Work

III. Fault Detection Approaches

VI. Conclusion

Authors

Figures

References

Citations

Keywords

Metrics

Abstract:
In these days, Wireless sensor network becomes integral part of the computer and communication technology. A huge number of sensor nodes can be deployed in the field of interest to monitor, detect the time critical events by which an emergency of the situation can be evaluated. All the sensor nodes in the network, may not perform well always and it may fail due to several reasons. Fault detection is considered as a vital factor in order to achieve the quality of service, to increase the bandwidth and the lifetime of the network. In this paper, we have classified various fault detection approaches used in wireless sensor networks.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date of Conference: 06-08 April 2017

INSPEC Accession Number: 17544901

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286633

ISBN Information:

Electronic **ISBN:**978-1-5090-3800-8

Print on Demand(PoD) **ISBN:**978-1-5090-3801-5

Publisher: IEEE

Conference Location: Chennai, India

I. Introduction

Abstract:
Vehicular Ad hoc Networking (VANET) is a wireless communication technology used in the vehicles to track the vehicle with the locations served by the Intelligent Transport System. The real-time traffic information is sought to the user and government officials by using these sorts of communication technologies. In order to be effective, the messaging system was implemented for VANET. Initially the process was very complex due to long message delay and the concept failed as there was a long message queue that led to heavy message loss. In order to avoid the queuing and time delay, the vehicle id is paired with location id to enhance the messaging service. This pairing helped the process to overcome and avoid co-operative messages which was used only for vehicle identity and its safety.

Published in: 2018 4th International Conference on Computing Communication and Automation (ICCCA)

Date of Conference: 14-15 December 2018

Date Added to IEEE *Xplore*: 29 July 2019

INSPEC Accession Number: 18868444

DOI: 10.1109/CCAA.2018.8777657

Publisher: IEEE

Conference Location: Greater Noida, India

▼ **ISBN Information:**

Electronic ISBN:978-1-5386-6947-1

USB ISBN:978-1-5386-6946-4

Print on Demand(PoD) ISBN:978-1-5386-6948-8

▸ **ISSN Information:**

4. Sivasankari, BhanuShri, Y. BevishJinila," Smart Waste Management Using WSN AND IOT,2017, Page No:6"

5. S.S. Navghane, M.S.Killendar, Dr. V.M. Rohokale,"IoT Based Smart Garbage and Waste Collection Bin,2016, Page No:7", International Journal Of Advance Research in electronics and communication engineering.

6. Sathish. A, Prakash. M,Jainulabudeen S.A.K, Sathishkumar R, 2017," Intellectual Trash Management using Internet of Things", Department of Computer Science Engineering, Panimalar Engineering College, Chennai, Tamilnadu, India. (ICCPEIC).

7. Keerthana B, Sonali M Raghavendran, Kalyani S, Suja P 5V.K.G. Kalaiselvi, 2017," INTERNET OF BINS-- Trash Management in India", Dept of Information Technology, Sri Sairam Engineering College, Chennai, India

applications" Advances in Wireless and Optical Communications (RTUWO).

12. Neetha, Sharma, S., Vaishnavi, V., &Bedhi, V (2017) Smart bin — An "Internet of Things approach to clean and safe public space", International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud.
13. Ravale, U., Khade, A., Patel, N., &Chaure, S. (2017), "Smart Trash: An Efficient Way for Monitoring Solid Waste Management", International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC).

14. Mustafa M.R, and Ku Azir K.N.F MATEC (2017) "SMART TRASH CAN", Web of Conferences.

15. AbhishekApte,DhirajGhadi, DevangDesai,Hiral (2015), "Waste Management System Using IOT", RaveshiyaNidhiSanghavi International Journal of Recent Trends in Engineering

Smart Waste Management Using WSN and IoT

[1]Sivasankari, [2]Bhanu Shri and [3]Y.Bevish Jinila

[1,2]Student, Department of IT, School of Computing, Sathyabama University
[3]Assistant Professor, Department of IT, School of Computing, Sathyabama University

***Abstract*— Rapid increase in population, has led to the improper waste management in cities resulting in increased pests and spreading of diseases. Nowadays, the Garbage Collecting Vehicle (GCV) collects the waste twice or thrice in a week. So, the problem is over flowing of wastages on the roads. Hence, to overcome this limitation, in this paper a scheme on smart waste management using Wireless Sensor Networks (WSN) and IoT (Internet of Things) is proposed. The garbage bins are deployed with sensors and are networked together using WSN. The sensors deployed in the garbage bins collect the data for every determined interval. Once the threshold is reached, it raises a request to the GCA (Garbage Collector Agent). This agent collects the requests of all the filled vehicles and communicate using IoT framework. The experimental simulation is done in proteus tool. A hardware prototype is developed for the proposed framework. Analysis of the proposed scheme provides better results in waste management.**

***Index Terms* – Wireless Sensor Networks, Smart Waste Management, Garbage Collection.**

I. INTRODUCTION

Solid waste is a general term that encloses all waste materials except unsure waste, liquid waste and region emissions. Solid wastes are often additional divided into two general groups: residential and business. The residential cluster refers to the waste that's labeled as 'garbage'. The rubbish category are often represented because the waste that's collected by the community services. The business cluster contains industrial and agricultural customers. These purchasers are far more spread and manufacture higher amounts of waste. The foremost necessary distinction between urban routes and business routes is that the waste management company traditionally, as a result of associate degree unloading trip needs to be create to a disposal location when every client.

II. RELATED WORKS

This section discusses about the existing approaches in the field of smart waste management.

Insung Hong et.al [1] has suggested that replacing SGS(Smart Garbage Sensor) instead of RFID garbage collecting system helps to improve their energy efficiency up to 16% and can reduce the food waste reduction .Inside the SGS they have installed SGBs(Smart Garbage Bins) to control the energy efficiency of the system.

Dario Bonino et.al [2] has suggested that it provides end-to-end security and privacy that is built upon dynamic federation smart city platform .Its benefits is that it has good dependability and has resilience on failure of a system over a particular month. It focuses on the collection of wastages and accomplishment of ontology method.

Alvaro Lozano Murciegoet.al [3] has suggested that to collect the dustbins that are been filled using a truck. The main advantage is that it reduces the fuel cost of the trucks rather than travelling a long distance it makes the path more simpler and easier to reach the dustbin using route optimization.

TheodorosAnagnostopoulos et.al [4] has suggested that it first starts with an assumption that the smart city must include the IoT base. It uses dynamic scheduling. It is based on the fact that the garbage will be collected only when it is fully filled or the maximum capacities of the dustbins are filled.

Rachael E. Marshall et.al [5] outlines that the smart waste management system in the high salaried countries and an developing countries.

Lilliana Abarca Guerrero et.al [6] outlines the fact that the developing countries undergo an prominent factor of affecting

A novel approach for mining temporal pattern database using greedy algorithm

Publisher: IEEE Cite This PDF

J. S. Vimali ; Sanjay Gupta ; Piyush Srivastava All Authors

4 Paper Citations

155 Full Text Views

Abstract

Document Sections

» Introduction

I. Temporal Database

II. Text Mining Process

III. Existing System

IV. Proposed System

Show Full Outline

Authors

Figures

References

Citations

Keywords

Abstract:
Consecutive example mining is a critical subfield in information mining and it's a troublesome issue since mining may require assembling or looking at expansive number of subsequence blends. Instructions to viably and productively mine interim based successions is testing issue. Three sort of interim based examples which utilize Greedy calculation, to find three kind of interim based successive examples. Time is inescapable of reality, and various social database approaches have been made to adjust to it. In realistic applications, facts can go over a couple times, and simply the general time span containing each one of the emphases may be known (consider, e.g., On January, John went to five social affairs of the Bioinformatics wander). While some brief social databases have defied, substances repeated at (known) incidental time, or single truths happened at fleetingly unverifiable time.

Published in: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

Date of Conference: 17-18 March 2017

Date Added to IEEE *Xplore*: 01 February 2018

ISBN Information:
Electronic ISBN:978-1-5090-3294-5
Print on Demand(PoD) ISBN:978-1-5090-3295-2

INSPEC Accession Number: 17558944

DOI: 10.1109/ICIIECS.2017.8275953

Publisher: IEEE

Conference Location: Coimbatore, India

▼ ISBN Information:

Electronic ISBN:978-1-5090-1277-0

DVD ISBN:978-1-5090-1276-3

Print on Demand(PoD) ISBN:978-1-5090-1278-7

Periodical:	Key Engineering Materials (Volume 759)
Pages:	76-80
DOI:	https://doi.org/10.4028/www.scientific.net/KEM.759.76
Citation:	Cite this paper
Online since:	January 2018
Authors:	Thankaraj Salammal Shyju*, Manidurai Paulraj
Keywords:	Anatase TiO_2, Nanocubes, Nanotubes, Solar Cell
Export:	RIS, BibTeX

MONOGRAPH ON COMPUTATIONAL INTELLIGENCE

FOR CLUSTER HEAD SELECTION IN WSN



by:
Dr. S. Emalda Roslin
Dr. N.M. Nandhitha
Dr. Rekha Chakravarthi
Ms. M.S. Sangeetha
Ms. Anita Daniel

A Monograph On Computational Intelligence For Cluster Head Selection In WSN

~~₹250.00~~ **₹220.00**

Author(s): Dr. S. Emalda Roslin, Dr. N.M. Nandhitha, Dr. Rekha Chakravarthi, Ms. M.S. Sangeetha and Ms. Anita Daniel

ISBN No.: 978-81-936671-8-7

Published Year: 2018

▼ ISBN Information:

Electronic ISBN:978-1-5090-1277-0

DVD ISBN:978-1-5090-1276-3

Print on Demand(PoD) ISBN:978-1-5090-1278-7

Springer Link

Published: 08 May 2018

A Hybrid Approach for Task Scheduling Using the Cuckoo and Harmony Search in Cloud Computing Environment

K. Pradeep ✉ & T. Prem Jacob

Wireless Personal Communications **101**, 2287–2311 (2018) | Cite this article

365 Accesses | **43** Citations | Metrics

Cite this article

Pradeep, K., Prem Jacob, T. A Hybrid Approach for Task Scheduling Using the Cuckoo and Harmony Search in Cloud Computing Environment. *Wireless Pers Commun* **101**, 2287–2311 (2018). https://doi.org/10.1007/s11277-018-5816-0

Download citation

Published
08 May 2018

Issue Date
August 2018

DOI
https://doi.org/10.1007/s11277-018-5816-0

Abstract

The proposal deals with designing of a novel framework for group decision support system (GDSS). The purpose of applying such a system is to broadcast the events of an earthquake occurrence based on group decisions by analyzing the reports from the sensor networks which are set in the deep sea. Recent advancements in the field of data analytics show that automatic decision making system is indispensable, exclusively in case of an emergency circumstance. Current frameworks are mostly based on single-user decision support systems. Well-established operational research techniques such as multi-criteria decision making techniques need to be integrated in order to bring out a successful decision support system. A comprehensive analysis of the challenges in the fields of information systems and machine learning frameworks is to be scrutinized. Such a study is critical for designing a framework based on platform independent automated multi-decision-based GDSS. This can efficiently support alternative types of goals and control protocols between its users.

Publisher Name	Print ISBN
Springer, Singapore	978-981-10-0766-8

A Survey on Collaborative Categorization using Fuzzy Logic for Improved User Suggestions

M. Anjani and V. Vijaya Kaveri

Sathyabama University, Chennai-600119, Tamil Nadu, India;
Anjani.m2008@gmail.com, vijirama.kaveri@gmail.com

Abstract

Background/Objectives: This paper presents a review on the usage of social tags that can be employed for recommending in social networks. It also focusses on pros and some algorithms as solutions to the problems. Recommender systems are the one used for providing information regarding the categories that belongs to the items which makes use of fuzzy clustering, in which the membership degree of the object is not available in the calculated categories. If calculated categories are available it will provide good quality recommendations to the user. **Methods/Analysis:** The method suggested in this work is to utilize fuzzy logic to classify the similar users from the clustered users based on their profiles and taste of viewing similar items. Fuzzy logic here plays a vital role in providing recommendation to the similar users once the classification is done. **Findings:** Hence, by surveying some papers on recommendation in social networking, we propose a solution by creating a dynamic system which would provision recommendations to the user to make the domain selection more effectively. The domains refers to books, movies, songs, which are purely dynamic. This methodology helps to display the most viewed item as the top most item and displays the item ratings in the webpage. The main approach used here is fuzzy logic with the detailed representation of features of objects and modeling user profiles. **Applications/Improvements:** This type of dynamic domain system will provide better results when compared to the conventional recommendation systems as it allows users to give comments on various products on different domains. This type of recommendation can be consumed by many social networking sites for gaining page rating and it can be further improved in future by combining the online dictionary to classify the terms used in user rating.

2nd International Conference on Sustainable Technologies in Building and Environment ICSTBE 2018 in February 2023.

Start Date	Fest Name	Fest Type	College Name & City
23 Jan 2018	2nd International Conference on Sustainable Technologies in Building and Environment ICSTBE 2018	Conference	Sathyabama University Chennai

Fasten your seatbelt with offers

Flat ₹1001 off on first flight

PRESENTATION SCHEDULE

SESSION 3 : GREEN TECHNOLOGY **SESSION CHAIR : Dr.N.K.Ambujam** **Co-Chair :Dr.Phani Madavi**

Date : 25.01.2018 **Time : 11.00 A.M – 12.30. P.M** **Venue : III** **Rapporteur : Ms.P.K.Raji**

S.No.	Participant Name	Title of Presentation
1.	Ahamed Fazeel Akram	Green building rating system
2.	Karan.R, Niyandha.S,Rajesh Khanna	Green Technology: an investigation on promoting green villages
3.	D.Mugundhan	Bridging green and brown foot prints – a sustainable challenge to envelope grey shade
4.	Nina Lazar, Chithra K	Green Building Rating Systems from the Perspective of Three Pillars of Sustainability through Building Lifecycle
5.	Selvendiran S G, Ebin Horrison S	A framework for evaluating the benefits of green building project using cost benefit analysis
6.	Andavan.S, Needhidasan.S, Elavenil.S	Implementing the concepts of green building by using eco-friendly materials for sustainable development – a study
7.	Anushruti, Mamatha.N	Smart &Sustainable materials: A step towards sustainability
8.	Ar.Vinushaa, Sukirtha, Surya Rajkumar	A study on centralized collaboration for sustainable development in AEC industry

Abstract:
Skin cancer cases have grown worldwide, and it is by far the most common of all cancers. It is caused by the extended exposure to harmful radiations from the sun. Skin cancer can be broadly classified into two categories as melanocytic and non-melanocytic (benign). Most basic way to detect the threat of skin is by visual investigation performed by a specialist dermatologist employing a set of specific clinical tools. This paper describes novel image processing approaches for skin cancer detection using dermoscopy images. The motive of this work is to investigate and prefer an algorithm for skin cancer determination that can group the lesions as malignant or benign melanoma in terms of accuracy, sensitivity and specificity. Initially, the image samples showing melanoma are obtained by using dermoscope and segmented. It is then given for feature extraction and the results are classified using Support Vector Machine (SVM), Artificial Neural Network (ANN) and Naive Bayes Classifiers.

Published in: 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)

Date of Conference: 27-28 July 2018

Date Added to IEEE *Xplore*: 28 February 2020

ISBN Information:
Electronic ISBN:978-1-5386-5995-3
Print on Demand(PoD) ISBN:978-1-5386-5996-0

INSPEC Accession Number: 19467707

DOI: 10.1109/ICRIEECE44171.2018.9008943

Publisher: IEEE

Conference Location: Bhubaneswar, India

Abstract:
The 2-Dimensional and 3-Dimensional navigation system makes the users to determine and it specifies the exact places in the map. This proposal deals with the mechanisms of the navigation technologies used to develop the 2-Dimensional and 3-Dimensional models. This makes the navigation easy and effective for the users to travel to their desired locations. The 2-D & 3-D navigation is much simple as the google maps, this requires the base map of the specified location to be designed. The 2-D navigation is much easier for the users to understand, since it is of uncomplicated format. When this comes to the 3-D navigation, this makes the users to experience that they travel in reality.

▼ ISBN Information:

Electronic ISBN:978-1-5386-0814-2

Print ISBN:978-1-5386-0813-5

CD:978-1-5386-0812-8

Print on Demand(PoD) ISBN:978-1-5386-0815-9

Handbook of Nanomaterials for Industrial Applications

Micro and Nano Technologies

2018, Pages 952-958



Chapter 52 - Impact of Engineered Nanomaterials for Environmental Industries

Vinita Vishwakarma

Centre for Nanoscience and Nanotechnology, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India

https://doi.org/10.1016/B978-0-12-813351-4.00054-7

By:
Dr. Rekha Chakravarthi
Dr. N.M. Nandhitha
Dr. S. Emalda Roslin
Ms. M.S. Sangeetha

INFRARED THERMOGRAPHY FOR CONDITION MONITORING OF ELECTRICAL EQUIPMENTS

Infrared Thermography For Condition Monitoring Of Electrical Equipments



Author(s): Dr. M.S. Sangeetha, Dr. N.M. Nandhitha, Dr. N.Selvarasu, Dr. S. Emalda Roslin and Dr. Rekha Chakravarthi

ISBN No.: 978-81-941110-1-6

Published Year: 2019

INTELLIGENT TECHNIQUES FOR CONGESTION CONTROL IN WSN



by:
Dr. Rekha Chakravarthi
Dr. N.M. Nandhitha
Dr. S. Emalda Roslin
Ms. M.S. Sangeetha

Author(s): Dr. Rekha Chakravarthi, Dr. N.M. Nandhitha, Dr. S. Emalda Roslin, Ms. M.S. Sangeetha and Mr. M. Mani Franklin

ISBN No.: 978-81-936671-6-3

Published Year: 2018

Nuclear power has been a big area of interest due to its efficiency. Having said that, safety concerns have been a key limiting factor in the adoption of nuclear reactors. In fact, even very developed countries have been considering the shutdown of these reactors. In such a scenario, the industrial community believes that the only way to increase the acceptance of nuclear energy is by building very good safety systems. Contribution from all domains of engineering is essential for developments in safety systems. To pave the way in this direction, a critical measurement system has been developed for sodium flow monitoring in Fast Breeder Reactor. This book describes the design and development of the measuring system which ensures the safety of the reactor. I believe that this book could be a good guide for anyone attempting to use electronics to build next-generation safety systems for any application that requires a faster dynamic response.

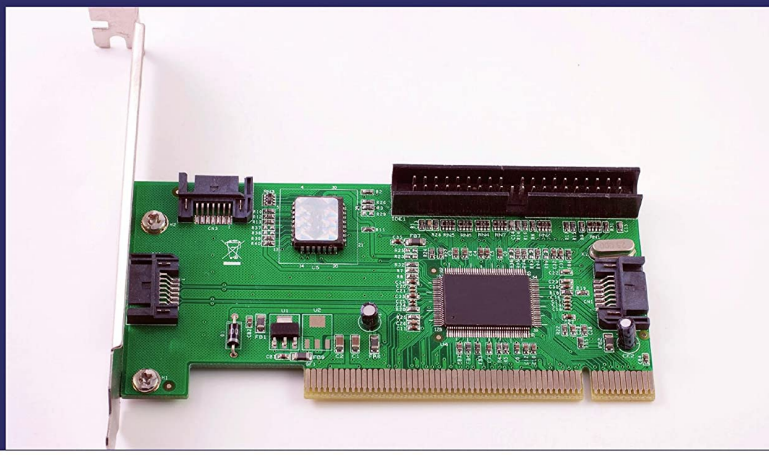


S.Poornapushpakala got her Bachelor's degree in Instrumentation and Control Engineering from Madras University in 1997. She completed her post graduation in Electronics and Control Engineering from Sathyabama University in 2004. She has completed her PhD in 2015 and has 15 publications & has successfully completed a funded project for IGCAR, India

978-620-2-07583-1

Critical Instru. for Flowmeters in FBTR

Suriyamoorthi



Poornapushpakala Suriyamoorthi

Critical Instrumentation for Flowmeters in FBTR

Eddy current flowmeter based Sodium flow Monitoring for safety of Nuclear Reactor

LAP LAMBERT Academic Publishing

Publisher : LAP LAMBERT Academic Publishing

Language : English

Paperback : 188 pages

ISBN-10 : 620207583X

ISBN-13 : 978-6202075831

Dynamic Modeling for Open- and Closed-loop Control of PMSG based WECS with Fuzzy Logic Controllers

WRITTEN BY

Anbarasi Jebaselvi Jeya Gnanaiah David and Meenakshi Veerappan

Submitted: May 18th, 2017 , Reviewed: November 24th, 2017 , Published: September 19th, 2018

DOI: 10.5772/intechopen.72693



FROM THE EDITED VOLUME

MATLAB

Edited by Ali Saghafinia

Book Details | Order Print

Copyrighted Material



Sakthi Prabha R

Fundamentals of Semiconductors

LAP LAMBERT Academic Publishing

Copyrighted Material

Publisher : LAP LAMBERT Academic Publishing

Language : English

Paperback : 52 pages

ISBN-10 : 6202023503

ISBN-13 : 978-6202023504

This monogram discuses novel techniques in image enhancement and segmentation of defects in radiographic weld image. To start with, a case study of Enhancement, Denoising, and segmentation techniques is analyzed and applied on the weld radiographic image. Implementation of these algorithms is done by MATLAB codes. The book is intended to the readers willing to expand their knowledge in the field of automated weld defect detection.



Dr.P. Chitra
Dr.N.M. Nandhitha
Dr.P. Kavipriya

Application of Image Processing for weld radiographic images

Dr.P.Chitra is currently working as an Associate Professor in Sethyabama University, Department of ECE. She had received her doctorate in Sathyabama University in Sep, 2014. She had more than 13 years of teaching experience and She had taught variety of courses for PG and UG. She had published more than 25 papers in reputed journals and conferences.

978-613-7-38034-5

LAMBERT Academic Publishing

Publisher : LAP LAMBERT Academic Publishing

Language : English

Paperback : 52 pages

ISBN-10 : 6137380343

ISBN-13 : 978-6137380345

Design Concepts of Low-Noise Amplifier for Radio Frequency Receivers

WRITTEN BY

Sumathi Manickam

Submitted: February 14th, 2018 , Reviewed: May 29th, 2018 , Published: November 5th, 2018

DOI: 10.5772/intechopen.79187



FROM THE EDITED VOLUME

RF Systems, Circuits and Components

Edited by Mamun Bin Ibne Reaz and Mohammad Arif Sobhan Bhuiyan

Book Details | Order Print

Design Concepts of Low-Noise Amplifier for Radio Frequency Receivers

WRITTEN BY

Sumathi Manickam

Submitted: February 14th, 2018 , Reviewed: May 29th, 2018 , Published: November 5th, 2018

DOI: 10.5772/intechopen.79187



FROM THE EDITED VOLUME

RF Systems, Circuits and Components

Edited by Mamun Bin Ibne Reaz and Mohammad Arif Sobhan Bhuiyan

Book Details | Order Print

INTELLIGENT TECHNIQUES FOR CONGESTION CONTROL IN WSN



by:
Dr. Rekha Chakravarthi
Dr. N.M. Nandhitha
Dr. S. Emalda Roslin
Ms. M.S. Sangeetha

Monograph On Intelligent Techniques For Congestion Control In WSN



Author(s): Dr. Rekha Chakravarthi, Dr. N.M. Nandhitha, Dr. S. Emalda Roslin, Ms. M.S. Sangeetha and Mr. M. Mani Franklin

ISBN No.: 978-81-936671-6-3

Published Year: 2018

MONOGRAPH ON COMPUTATIONAL INTELLIGENCE

FOR CLUSTER HEAD SELECTION IN WSN



by:
Dr. S. Emalda Roslin
Dr. N.M. Nandhitha
Dr. Rekha Chakravarthi
Ms. M.S. Sangeetha
Ms. Anita Daniel

A Monograph On Computational Intelligence For Cluster Head Selection In WSN



Author(s): Dr. S. Emalda Roslin, Dr. N.M. Nandhitha, Dr. Rekha Chakravarthi, Ms. M.S. Sangeetha and Ms. Anita Daniel

ISBN No.: 978-81-936671-8-7

Published Year: 2018

By:
Dr. Rekha Chakravarthi
Dr. N.M. Nandhitha
Dr. S. Emalda Roslin
Ms. M.S. Sangeetha

INFRARED THERMOGRAPHY FOR CONDITION MONITORING OF ELECTRICAL EQUIPMENTS

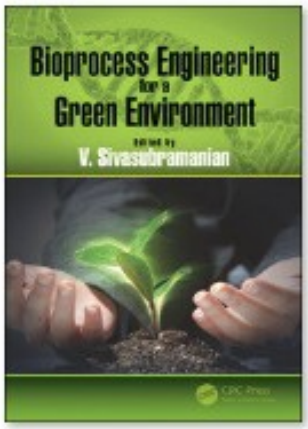
Infrared Thermography For Condition Monitoring Of Electrical Equipments

~~₹300.00~~ **₹250.00**

Author(s): Dr. M.S. Sangeetha, Dr. N.M. Nandhitha, Dr. N.Selvarasu, Dr. S. Emalda Roslin and Dr. Rekha Chakravarthi

ISBN No.: 978-81-941110-1-6

Published Year: 2019



Book

Bioprocess Engineering for a Green Environment

Edited By V. Sivasubramanian

Edition	1st Edition
First Published	2018
eBook Published	4 May 2018
Pub. Location	Boca Raton
Imprint	CRC Press
DOI	https://doi.org/10.1201/b22021
Pages	438
eBook ISBN	9781315232379
Subjects	Bioscience, Engineering & Technology

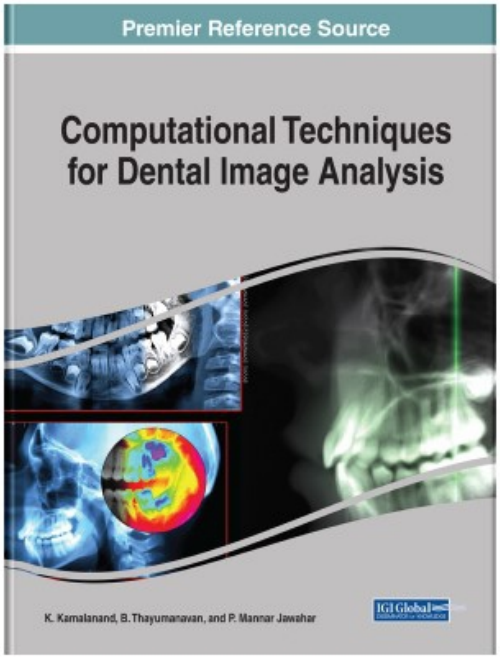
Chapter 14 | 20 pages

Characterization and Optimization Studies on Hydroxyapatite Bioceramic Powder from Waste Eggshells

By A. Annam Renita, V. Sivasubramanian, P. Senthil Kumar

Abstract ^

Hydroxyapatite is a pale yellow, moldable, semivitreous solid that exhibits ceramic characteristics. It is similar to the apatite found in human bones and teeth. Hence, it finds application in biomedical and -chemical fields as, for example, a bone substitute, a dental filler, catalyst, in tissue engineering, and in drug delivery arenas. It can be synthesized from synthetic and biological raw materials. Preparing hydroxyapatite from synthetic sources is not cost effective, as costly and pure precursors are required. Because there is a need for clean technologies and an inclination for substituting synthetics with green chemistry, hydroxyapatite synthesized from biomaterials has been investigated. Biological raw materials containing calcium and magnesium sources can be used as starting materials for hydroxyapatite synthesis. Sources include fish bone, bovine bone, oyster shells, cockle shells, snail shells, chicken eggshells, crab shells, and mussel shells. This chapter outlines various synthesis methods of hydroxyapatite, applications of hydroxyapatite, an experimental investigation of production of hydroxyapatite by the precipitation method, optimization of reaction conditions, and characterizations of hydroxyapatite. Experimental investigations have synthesized hydroxyapatite from waste eggshells via the wet precipitation method. Hydroxyapatite reaction conditions were optimized via response surface methodology, and the optimum conditions for the production of hydroxyapatite from eggshells was found to be 2 Ca/P ratio, 600 rpm agitation speed, and 30 minutes time. The predicted R-squared value of 0.957 is in reasonable agreement with the adjusted R-squared value of 0.985. Hydroxyapatite was characterized by FTIR, XRD, and SEM, which confirmed the presence of CO_3^{2-} groups and PO_4^{3-} groups with hexagonal dipyramidal structure.



IGI Global, 2019. — 359 p. — ISBN: 9781522562436

With the technology innovations dentistry has witnessed in all its branches over the past three decades, the need for more precise diagnostic tools and advanced imaging methods has become mandatory across the industry. Recent advancements to imaging systems are playing an important role in efficient diagnoses, treatments, and surgeries.Computational Techniques for Dental Image Analysis provides innovative insights into computerized methods for automated analysis. The research presented within this publication explores pattern recognition, oral pathologies, and diagnostic processing. It is designed for dentists, professionals, medical educators, medical imaging technicians, researchers, oral surgeons, and students, and covers topics centered on easier assessment of complex cranio-facial tissues and the accurate diagnosis of various lesions at early stages.

Dental Imaging and Analysis

Digital Image Enhancement Techniques for Dental Radiographs: A Support to Clinicians
Thresholding Techniques for Dental Radiographic Images: A Comparative Study
Image Segmentation Using Contour Models: Dental X-Ray Image Segmentation and Analysis
Dental Image Segmentation Using Clustering Techniques and Level Set Methods
Jaya Algorithm-Assisted Evaluation of Tooth Elements Using Digital Bitewing Radiography Images
Teeth and Landmarks Detection and Classification Based on Deep Neural Networks
Dental Cone Beam Computed Tomography for Trabecular Bone Quality Analysis in Maxilla and Mandible

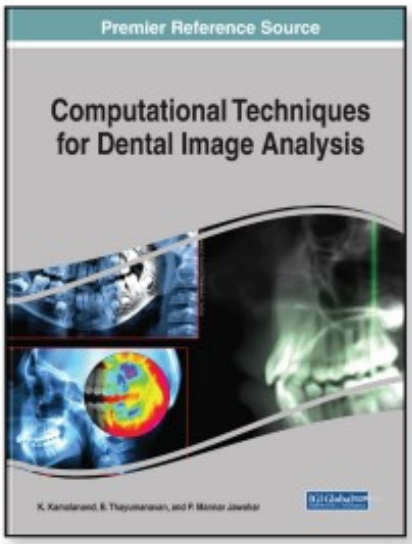
Dental Materials, Mechanics, and Instrumentation

Comparison and Analysis of Dental Imaging Techniques
Thermal Analysis of Artificial Dental Materials Using Numerical Simulation
Recent Trends in 3D Printing of Dental Models: Rapid Prototyping in Dental Implants
Phytochemicals: Their Therapeutic Potential Against Dental Caries

IGI Global

Publisher of Peer-Reviewed, Timely, and Innovative Research Content

Computational Techniques for Dental Image Analysis



Part of the Advances in Medical Technologies and Clinical Practice Book Series

K. Kamalanand (Anna University, India), B. Thayumanavan (Sathyabama University Dental College and Hospital, India) and P. Mannar Jawahar (Anna University, India)

Description:

With the technology innovations dentistry has witnessed in all its branches over the past three decades, the need for more precise diagnostic tools and advanced imaging methods has become mandatory across the industry. Recent advancements to imaging systems are playing an important role in efficient diagnoses, treatments, and surgeries.

Computational Techniques for Dental Image Analysis provides innovative insights into computerized methods for automated analysis. The research presented within this publication explores pattern recognition, oral pathologies, and diagnostic processing. It is designed for dentists, professionals, medical educators, medical imaging technicians, researchers, oral surgeons, and students, and covers topics centered on easier assessment of complex cranio-facial tissues and the accurate diagnosis of various lesions at early stages.

ISBN: 9781522562436 **Release Date:** October, 2018 **Copyright:** 2019 **Pages:** 280

Topics Covered

ORGANIC CHEMISTRY MANUAL



Dr.P.Shanmuga Pandiyan

Urün Bilgileri

Yayıncı : Independently published

Dil : İngilizce

ISBN-10 : 1792649274

ISBN-13 : 978-1792649271

RSIRJLE GWALIOR Research Scholar - An International Refereed Journal of Literary Explorations

ISSN 2320 – 6101
www.researchscholar.co.in

Research Scholar
An International Refereed e-Journal of Literary Explorations

REMEDIAL TEACHING THROUGH REFINED TEACHERS AND TECHNOLOGY

Dr. S.Mercy Gnana Gandhi
Professor in English-ELT
Sathyabama University
Chennai

ABSTRACT

Teacher motivation and confidence are the direct determinants in student classroom learning. Successful technology integration requires teacher skill, training, motivation and access to applicable tools. Students rarely respond positively to technology exposure when there is a lack of endorsement by the teacher. Teachers are quick to overcome factors such as anxiety in using new technology, but show to change their perception on perceived benefits. With respect to learning theory, we are concerned with processes of learning and conditions that promote successful language learning. Teachers may develop their own teaching procedures and theory of learning. They may constantly revise, vary, and modify teaching/learning procedures on the basis of the performance of the learners. Theory does not dictate a particular set of Teaching Techniques and activities. What links the theory with practice is design.

Key words: Remedial Teaching, Teachers, Technology



Mercy Gnana Gandhi

Professional Development in English Language Teaching,

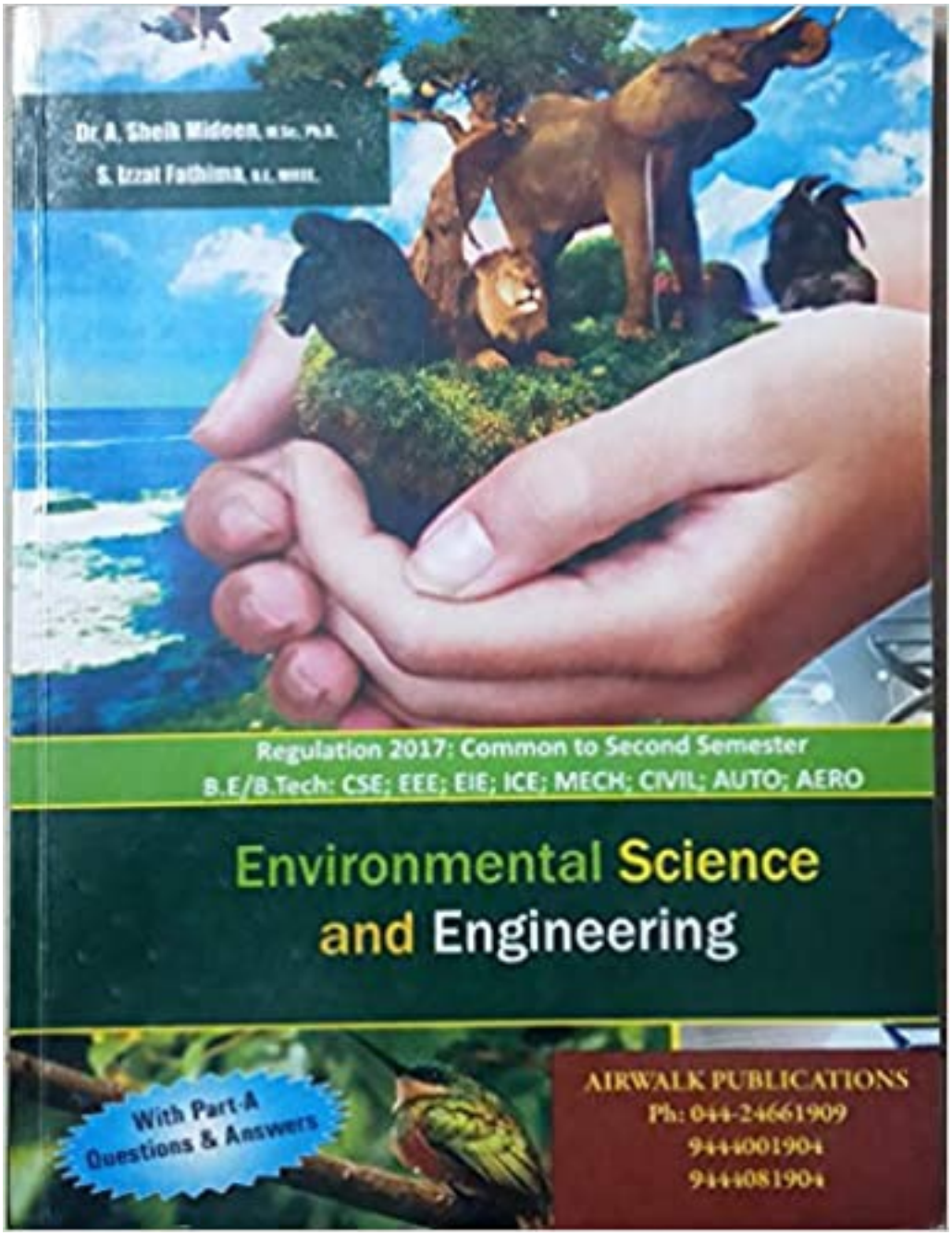
Authors S MERCY GNANA GANDHI

Publication date 2017

Volume 1

Pages Aakash Publications

Publisher ISBN 978-93-83


Dr. A. Sheik Mideen, M.Sc., Ph.D.
S. Izzat Fathima, B.E., M.E.
Regulation 2017: Common to Second Semester
B.E/B.Tech: CSE; EEE; EIE; ICE; MECH; CIVIL; AUTO; AERO
Environmental Science and Engineering
With Part-A Questions & Answers
AIRWALK PUBLICATIONS
Ph: 044-24661909
9444001904
9444081904

Product details

Publisher : AIRWALK PUBLICATIONS

Language : English

ISBN-10 : 9384893838

ISBN-13 : 978-9384893835

KINEMATICS OF MACHINERY

With Part -A Questions & Answers
and University Solved Papers





FOR III SEMESTER B.E
MECHANICAL ENGINEERING STUDENTS

Rs.350/-
After discount 20%
Rs.280/-



As Per Anna University
New Revised Syllabus



Dr. S. Ramachandran
M. Purusothaman



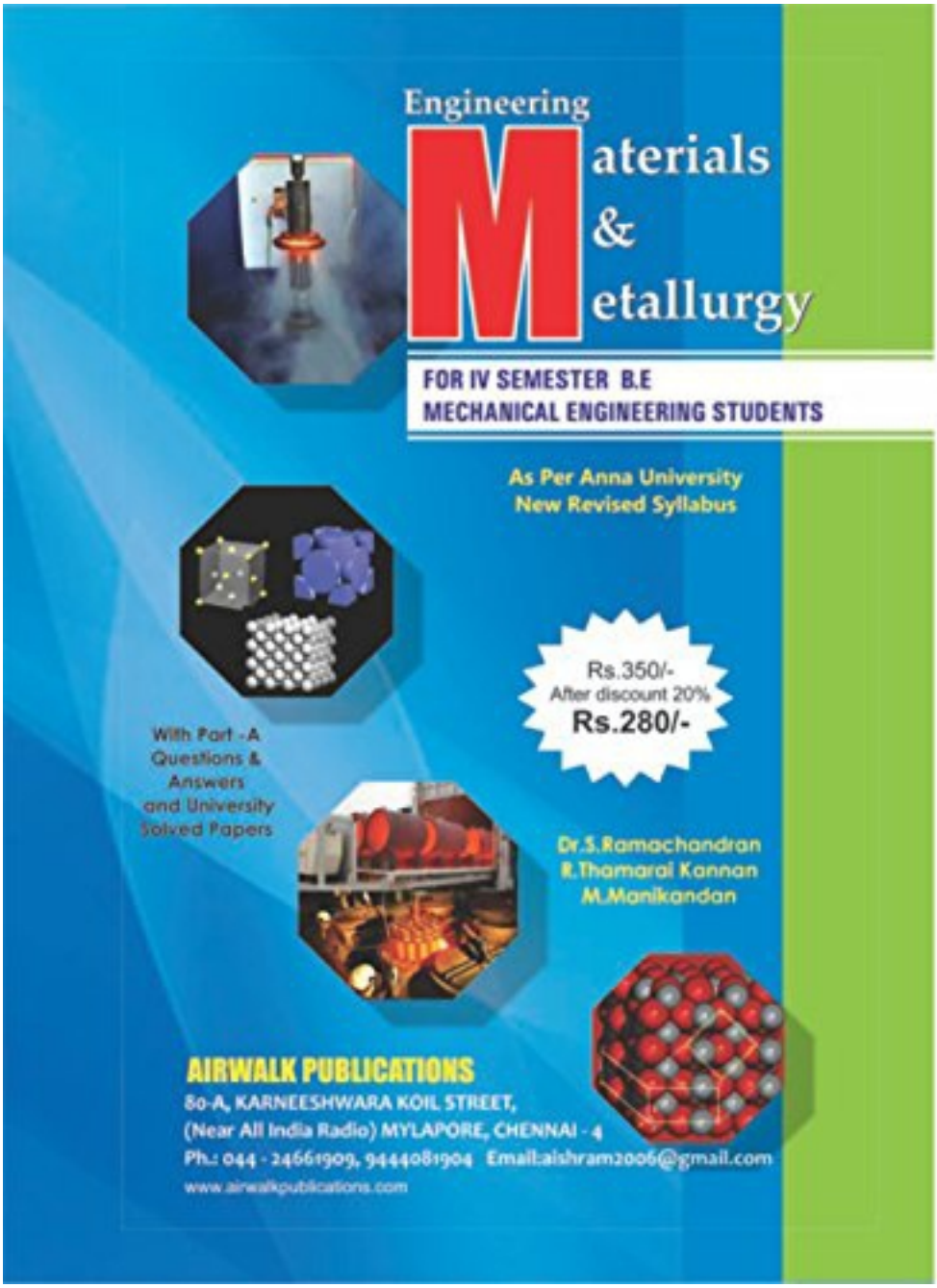




AIRWALK PUBLICATIONS
80, KARNEESHWARA KOIL STREET,
(Near All India Radio) MYLAPORE, CHENNAI - 4
Ph.: 044 - 24661909, 9444081904
Email: aishram2006@gmail.com www.airwalkpublications.com

ISBN : 978-81-924031-6-8




Engineering
Materials
&
Metallurgy
FOR IV SEMESTER B.E
MECHANICAL ENGINEERING STUDENTS
As Per Anna University
New Revised Syllabus
Rs.350/-
After discount 20%
Rs.280/-
With Part - A
Questions &
Answers
and University
Solved Papers
Dr.S.Ramachandran
R.Thamarai Kannan
M.Manikandan
AIRWALK PUBLICATIONS
80-A, KARNEESHWARA KOIL STREET,
(Near All India Radio) MYLAPORE, CHENNAI - 4
Ph.: 044 - 24661909, 9444081904 Email:aishram2006@gmail.com
www.airwalkpublications.com

ISBN : 978-81-924031-2-0



IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Synthesis and Characterization of bioactive composite coating on Titanium by PVD for biomedical application

Karthik Alagarsamy[1], Vinita Vishwakarma[1] and Gobi Saravanan Kaliaraj[1]

Published under licence by IOP Publishing Ltd

IOP Conference Series: Materials Science and Engineering, Volume 561, First International Conference on Materials Science and Manufacturing Technology 12–13 April 2019, Hotel Aloft, Coimbatore, Tamil Nadu, India

Citation Karthik Alagarsamy *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **561** 012027

DOI 10.1088/1757-899X/561/1/012027


materialstoday:
PROCEEDINGS

https://doi.org/10.1016/j.matpr.2016.04.041

Get rights and content

Abstract

The Zinc Oxide (ZnO) thin film deposited on n-type silicon (Si) by radio frequency magnetron sputtering at room temperature. During the Radio Frequency (RF) power, argon/oxygen gas pressure at different Standard Cubic Centimeters per Minute (SCCM), substrate temperature was varied, optimized for crystal structure, surface morphology, roughness, and optical electrical studies. The deposited films were examined crystalline structures and investigated by X-Ray Diffraction (XRD), surface morphology by Field Emission Scanning Electron Microscopy (FESEM) and Atomic Force Microscopy (AFM), optical studies by UV-Vis, electrical studies by IV characteristic. The film post deposition annealed was improving the quality of the film properties

Electric properties of ZnO thin films by RF Magnetron sputtering technique ☆

K. Viswanathan, T.S. Shyju, D. Ramachandran, G. Pradhaban

Centre for Nanoscience and Nanotechnology, Sathyabama University,Chennai-600119, India

Available online 2 June 2016, Version of Record 2 June 2016.

A New Approach to Grow A-TiO_2 Nanocubes

608



Abstract:

The vertically aligned TiO_2 Nanotubes (TiNTs) extracts electrons from an absorber and also helps in its transport in perovskite and dye sensitized solar cells (DSSCs) solar cells. Thus electron transporting layer plays a very important role in photon to electron conversion. Electrochemical anodization is been used widely to grow TiNTs for solar cell applications; due to its low cost, flexibility to vary pore diameter and tube length. We observed that, TiNTs maintained its tubular array only for a set time period. On increasing the growth time they begin to take the form of nanocrystals with {001} facets. Formation of these titanium nanocrystals (TiNcs) was clearly observed through field emission scanning electron microscope (FESEM) and Transmission electron microscopy (TEM). Thus TiO_2 nanostructures can be tuned by varying anodization time. More importantly, optimization of the reaction process led to the growth of more orderly, crystalline anatase TiNTs/TiNcs over Ti metal foil substrate. The crystal structure and surface morphology of the prepared thin film samples were studied using X-ray diffraction (XRD) technique and scanning and transmission electron microscopes (TEM). XRD confirmed the anatase phase of as grown TiO_2 with (101) as major intensity preferred orientation. Lattice parameters calculated were found to be a= 3.77-to-3.82 and c= 9.42-to-9.58 for grown TiO_2.

Functional Materials and Application

ISBN(softcover): 978-3-0357-1219-3 ISBN(eBook): 978-3-0357-3219-1

Table of Contents

Preface, Conference Committees

Chapter 1: Properties of the Structural Materials

Waste and Discharge Sand Use of Foundry for the Manufacture of Concrete Tiles
A.V. dos Santos, M.C. Marchetti, A.R. de Souza, D.J.C. Silva, L.P. Godoy, V.A. Casarin and M.A.G. Marchetti 3

The Mechanical and Physical Properties of Construction and Demolition Waste - Epoxy Composites
M. Hyvärinen and T. Kärki 9

Fabrication of Al-Si-Mg Functional Composites Engine Cylinder and its Bench Test
X.D. Lin, C.M. Liu and J. Sun 15

Effect of Treatment of Fly Ash on Mechanical Properties of Polypropylene
P. Doddipatla and S. Agrawal 20

Refinement Mechanism of Solidification Structure of Au-20Sn Eutectic Alloy by Different Solidification Techniques
Y. Mao, J.X. Guo and S.Y. Xu 24

Investigation of Flux-Less Soldering of Metal Matrix Composite and Ceramics by Use of Zn Solder
R. Koleňák and I. Kostolný 29

Characterization Investigations of a Al-5 wt.% Si/2 wt% (NbB_2, NbC) Hybrid Composite Fabricated via Mechanically Alloying and Sequentially Milling (Mechanical Alloying + Cryomilling)
M.L. Öveçoğlu, E. Tekoğlu, S. Mertdinç and D. Ağaoğulları 35

Chapter 2: Surface Engineering, Thin Films and Special Coatings

Computational Prediction of Trimethylgallium Adsorption on Si(100)(2×1) in Atomic Layer Deposition
S. Sukhasena and P.P. Pansila 43

Effect of Electron Bombardment on Polyimide Film
D.P. Engelhart, E. Plis, R. Cooper, S. Humagain, A. Koch, M. Brunetti, S. Greenbaum and R. Hoffmann 48

A Novel Preparation Method of $CH_3NH_3PbI_3$ Thin Films for Perovskite Solar Cell
W.L. Zhao 54

Study on the Structural Analysis of Al_2O_3 Ceramic-Lined Steel Pipe Prepared with the Gravitational Separation SHS Method
X.H. Hou, J.K. Yu, X. Yang and Z.Y. Liu 59

Investigation of Changing of Chemical and Phase Composition of the Friction Surface Fe-Mn-C-B-Si-Ni-Cr
M. Pashechko, K. Lenik, K. Dziedzic and M. Barszcz 65

Chapter 3: Nanomaterials and Nanotechnologies

M. Pashechko, K. Lenik, K. Dziedzic and M. Barszcz 65

Chapter 3: Nanomaterials and Nanotechnologies

Effect of Lattice Mismatch Strain Grading on the Electromechanical Behavior of Functionally Graded Quantum Dots
P.L. Bishay, B. Sampat, J. Sladek, E. Pan and V. Sladek 71

A New Approach to Grow A-TiO_2 Nanocubes
T.S. Shyju and M. Paulraj 76

Effect of ZnO Concentration on the Diameter of Electrospun Fibers from Poly(Vinyl Alcohol) Composited with ZnO Nanoparticles
T. Jamnongkan, S. Kaewpirom, A. Wattanakornsiri and R. Mongkholrattanasit 81

Chapter 7 - Green Synthesis of Metal Nanoparticles Using Seaweed Polysaccharides

Gobi S. Kaliaraj [*], Bharathiraja Subramaniyan [**], Panchanathan Manivasagan [†], Se-Kwon Kim [†]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/B978-0-12-809816-5.00007-4

Get rights and content

S. Gowtham Kumar
V. K. Langeswaran
Editors

Cancer versus Nutraceuticals



Cancer Etiology, Diagnosis and Treatments

Nova Biomedical

NOVA

ISBN: 978-1-53612-529-0

Categories: **2017**, **Cancer Etiology, Diagnosis and Treatments**, **Medicine and Health**, **Oncology**

Tags: **9781536125290**, **9781536125313**, **oncology**

Table of Contents

Preface

Chapter 1. Nutraceutical Therapeutics against Cancer (pp. 1-42)
V. K. Langeswaran

Chapter 2. The Role of Nutraceuticals in Cancer Prevention (pp. 43-64)
S. Perumal

Chapter 3. Cancers and Curcumin as a Nutraceutical Remedy (pp. 65-96)
Solai Priya



Jerrine Joseph

A review on Selected Medicinal plants

Taxonomy, Distribution, Phytochemistry And Biomedical Activities Of Selected Indian Medicinal Plants

LAP LAMBERT Academic Publishing

A review on Selected Medicinal plants : Taxonomy, Distribution, Phytochemistry And Biomedical Activities Of Selected Indian Medicinal Plants

Jerrine Joseph

Published by LAP LAMBERT Academic Publishing (2018)
ISBN 10: 6139877911 / ISBN 13: 9786139877911
New / Taschenbuch / Quantity: 1

Title	Fermentation Technology: Laboratory Manual
Authors	Radhakrishnan Manikkam Balagurunathan Ramasamy, Gopikrishnan Venugopal Premalata Pati, Jerrine Joseph Shanmugasundaram Thangavel
Publisher	LAP LAMBERT Academic Publishing, 2018

Understanding ENZYMES

An Introductory Text

Aditya Arya Amit Kumar Jayanti Jha



First Edition

ASIN : B07B3G3NB4

Publisher : Drawing Pin Publishing

Paperback : 220 pages

ISBN-10 : 8193674006

ISBN-13 : 978-8193674000

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-0901-5

CD:978-1-5090-0900-8

Print on Demand(PoD) ISBN:978-1-5090-0902-2

Publisher: IEEE

Conference Location: Melmaruvathur, India

Wireless expounding of voltage, current, temperature, gas in transformer

Publisher: IEEE

Cite This

PDF

M. Pushpavalli ; P. Sivagami **All Authors**

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290465

ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 International Conference...

Voltage sag reduction and power quality improvement using DVR

Publisher: IEEE

Cite This PDF

S. Sathish Kishore ; Sushant Kumar Sinha ; P. Abirami ; Merin Lizbeth George **All Authors**

Conferences > 2017 International Conference...

Tracking the calorie consumption and location using existing sensors and mobile applications

Publisher: IEEE

Cite This PDF

Venishetti RajaPavan ; Bulusu Sri Sai Suhas ; Ankayar Kanni **All Authors**

Date Added to IEEE *Xplore*: 05 October 2017

DOI: 10.1109/I-SMAC.2017.8058242

ISBN Information:

Publisher: IEEE

Conference Location: Palladam, India

J Aravind Kumar

D Joshua Amarnath

S Anuradha Jabasingh

S Sathish

DOI: 10.5897/AJB2016.15650

Keywords: Naphthalene, Borassus flabellifer, thermo-chemical, activated carbon, adsorption

Abstract

The present research describes the removal of naphthalene from aqueous solution using one of the simplest agricultural wastes, *Borassus flabellifer* Shell activated carbon (BFS-AC) by adsorption. Adsorption was chosen due to its cost-effectiveness. The effect of operating parameters such as pH, time, initial concentration of naphthalene, BFS-AC dosage and temperature on the percent removal of naphthalene was determined. The optimum conditions obtained were pH of 7, contact time of 12 h for an optimum concentration of 200 mg/L with the addition of 8 g at a temperature optimum at 40°C attaining the maximum percent removal. The adsorption kinetics supported the Freundlich model with R2 value

Chapter

The role of monoterpenes in cancer: A natural chemotherapy

January 2017

Project: Geraniol and Oral carcinogensis

Authors:



Madan Kumar
Sathyabama Institute of Science and Tec...

Download citation

Copy link

Abstract

Cancer is a group of diseases characterized by abnormal cell growth with the potential to spread to other parts of the body. Cancer's have become a major concern across the globe. Every year 4548 per 1,000,000 people are being reported by cancer, according to a survey in 2008-2012. Isolation and purification of anti-cancer agents from plants and other available natural sources has become a common phenomenon and successful indeed. Terpenoids are a wide range of phytochemical active compounds which include volatile, non-volatile, aromatic as well as nonaromatic constituents. Monoterpenes are a class of terpenes and nonnutritive dietary products which are produced mainly by the plants.Studies on these monoterpenes have revealed that these compounds have a variety of beneficial effects on human health such as antibacterial, antifungal, antiviral, antioxidant antitumor and anti-inflammatory effect. Also, these nutraceutical compounds can be isolated from plant sources showing notable beneficial effects against various cancer types.

In: Cancer versus Nutraceuticals
Editors: S. Gowtham Kumar et al.
ISBN: 978-1-53612-529-0
© 2017 Nova Science Publishers, Inc.

Chapter 5

THE ROLE OF MONOTERPENES IN CANCER: A NATURAL CHEMOTHERAPY

A. Madankumar, PhD*
Cancer Biology Lab, Molecular and Nanomedicine Research Unit,
Sathyabama University, Chennai,
Tamil Nadu, India

ABSTRACT

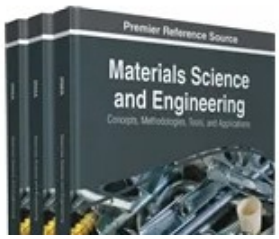
Cancer is a group of diseases characterized by abnormal cell growth with the potential to spread to other parts of the body. Cancer's have become a major concern across the globe. Every year 4548 per 1,000,000 people are being reported by cancer, according to a survey in 2008-2012. Isolation and purification of anti-cancer agents from plants and other available natural sources has become a common phenomenon and successful indeed. Terpenoids are a wide range of phytochemical active compounds which include volatile, non-volatile, aromatic as well as non-aromatic constituents. Monoterpenes are a class of terpenes and non-nutritive dietary products which are produced mainly by the plants.

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Login/Create Account Language: Eng

Search title, author, ISBN...

Books Journals e-Collections Open Access Publish with Us Resources Catalogs About



The Impact of Nanotechnology on Environment

Rathinasabapathi Jayavarthanan (Sathyabama University, India), Anima Nanda (Sathyabama University, India), and M. Amin Bhat (Sathyabama University, India)

Source Title: Materials Science and Engineering: Concepts, Methodologies, Tools, and Applications

Copyright: © 2017 | Pages: 31

DOI: 10.4018/978-1-5225-1798-6.ch068

Chapter 64
Teaching "Design-for-Corrosion" to Engineering Undergraduates: A Case Study of Novel Ni-B Coatings for High Wear and Corrosive Applications 1578
Ramazan Kahraman, Qatar University, Qatar
R. A. Shakoor, Qatar University, Qatar

Chapter 65
Risks and Preventive Measures of Nanotechnology 1605
Waqas Anwar, Mirpur University of Science and Technology, Pakistan
Anwar Khitab, Mirpur University of Science and Technology, Pakistan

Chapter 66
Quasi-SMILES for Nano-QSAR Prediction of Toxic Effect of Al2O3 Nanoparticles 1624
Alla P. Toropova, IRCCS, Istituto di Ricerche Farmacologiche Mario Negri, Italy
P. Ganga Raju Achary, Institute of Technical Education and Research (ITER), Siksha 'O'Anusandhan University, India
Andrey A. Toropov, IRCCS, Istituto di Ricerche Farmacologiche Mario Negri, Italy

Chapter 67
Linking Materials Science and Engineering Curriculum to Design and Manufacturing Challenges of the Automotive Industry 1636
Fugen Daver, RMIT University, Australia
Roger Hadgraft, CQ University, Australia

Chapter 68
The Impact of Nanotechnology on Environment 1659
Rathinasabapathi Jayavarthanan, Sathyabama University, India
Anima Nanda, Sathyabama University, India
M. Amin Bhat, Sathyabama University, India

Conferences > 2017 International Conference...

THD analysis between 5LT^{2}C and 7LT^{2}C inverter

Publisher: IEEE

Cite This

PDF

Merin Lizbeth George ; P. Abirami **All Authors**

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290446

ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 International Conference...

Stock prediction and analysis using intermittent training data with artificial neural networks

Publisher: IEEE

Cite This PDF

N. Srinivasan ; C. Lakshmi **All Authors**

Date Added to IEEE *Xplore*: 01 February 2018

DOI: 10.1109/ICIIECS.2017.8276082

ISBN Information:

Electronic ISBN:978-1-5090-3294-5

Print on Demand(PoD) ISBN:978-1-5090-3295-2

Publisher: IEEE

Conference Location: Coimbatore, India

Date of Conference: 14-16 December 2017

Date Added to IEEE *Xplore*: 08 November 2018

ISBN Information:

Electronic ISBN:978-1-5090-6621-6

CD:978-1-5090-6619-3

Print on Demand(PoD) ISBN:978-1-5090-6622-3

ISSN Information:

Electronic ISSN: 2473-943X

Print on Demand(PoD) ISSN: 2471-7851

INSPEC Accession Number: 18248602

DOI: 10.1109/ICCIC.2017.8524313

Publisher: IEEE

Conference Location: Coimbatore, India

Conferences > 2017 IEEE International Confe...

Simulation and Implementation of Two-Inductor PFC Boost Converter Using Single Pole Double Throw Switch (SPDT) Relay

Publisher: IEEE

Cite This PDF

G.T. Sundar Rajan ; Barnabas Paul Glady J. ; A. Ramesh Babu **All Authors**

Conferences > 2017 International Conference...

Sharpening of edges in radiographic images using

Publisher: IEEE

Cite This

PDF

P. Chitra ; Jintu Manuel **All Authors**

Date Added to IEEE *Xplore*: 16 October 2017

DOI: 10.1109/ICISC.2017.8068617

ISBN Information:

Electronic ISBN:978-1-5090-4715-4

Print on Demand(PoD) ISBN:978-1-5090-4716-1

Publisher: IEEE

Conference Location: Coimbatore, India

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-3751-3

DVD ISBN:978-1-5090-3750-6

Print on Demand(PoD) ISBN:978-1-5090-3752-0

Publisher: IEEE

Conference Location: Kollam, India

Review on design of less complex, high speed Viterbi decoders for accurate extraction of message bits

Publisher: IEEE

Cite This

PDF

B.N. HariChandana ; Yenduri. Bhagyasri ; V. Balamurugan ; N. M. Nandhitha **All Authors**

193
Full
Text Views

▼ ISBN Information:
Electronic ISBN:978-1-5090-4556-3
Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Conferences > 2017 International Conference...

Research and development of a mobile based women safety application with real-time database and data-stream network

Publisher: IEEE

Cite This PDF

Dantu Sai Prashanth ; Gautam Patel ; B. Bharathi All Authors

Published in: 2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)

Date of Conference: 20-21 April 2017

Date Added to IEEE *Xplore*: 19 October 2017

ISBN Information:

Electronic ISBN:978-1-5090-4967-7

Print on Demand(PoD) ISBN:978-1-5090-4968-4

INSPEC Accession Number: 17279830

DOI: 10.1109/ICCPCT.2017.8074261

Publisher: IEEE

Conference Location: Kollam, India

Conferences > 2017 International Conference...

Research and development of a mobile based women safety application with real-time database and data-stream network

Publisher: IEEE

Cite This PDF

Dantu Sai Prashanth ; Gautam Patel ; B. Bharathi All Authors

Published in: 2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)

Date of Conference: 20-21 April 2017

Date Added to IEEE *Xplore*: 19 October 2017

ISBN Information:

Electronic ISBN:978-1-5090-4967-7

Print on Demand(PoD) ISBN:978-1-5090-4968-4

INSPEC Accession Number: 17279830

DOI: 10.1109/ICCPCT.2017.8074261

Publisher: IEEE

Conference Location: Kollam, India

Conferences > 2017 International Conference...

Reduction of harmonic distortion in micro grids and distribution systems with photovoltaic generators

Publisher: IEEE

Cite This PDF

K. Prasanna Kumar ; R. Vanitha ; S. Sai Hemesh All Authors

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290466

▼ ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 Third International Conf...

Radio frequency identification technology for advanced library management

Publisher: IEEE

Cite This PDF

V. Senthil Nayagam ; Saikrishna Vedullapalli ; Mamidi Saikrishna ; R. S. Sivakumar All Authors

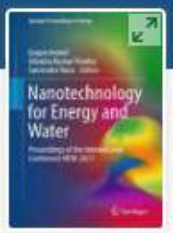
Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261401

▸ ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India



International Conference on Nano for Energy and Water
↳ ICNEW 2017: **Nanotechnology for Energy and Water** pp 131–139 | Cite as

Qualitative and Quantitative Analysis of Lipid Derived via Cell Disruption Techniques from *Chlorella vulgaris* for Biodiesel Production

N. Nirmala ✉, S. S. Dawn & S. Shiyam Rengith

Conference paper | First Online: 27 September 2017

646 Accesses

Part of the Springer Proceedings in Energy book series (SPE)

DOI
https://doi.org/10.1007/978-3-319-63085-4_18

Published
27 September 2017

Publisher Name
Springer, Cham

Print ISBN
978-3-319-63084-7

Online ISBN
978-3-319-63085-4

eBook Packages
Energy
Energy (R0)

Conferences > 2017 Third International Conf...

Quadcopter based technology for an emergency healthcare

Publisher: IEEE

Cite This PDF

A. Josephin Arockia Dhivya ; J. Premkumar All Authors

Date Added to IEEE *Xplore*: 26 October 2017

DOI: 10.1109/ICBSII.2017.8082284

▼ ISBN Information:

Electronic ISBN:978-1-5090-4980-6

Print on Demand(PoD) ISBN:978-1-5090-4981-3

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Providing cluster categorization of heuristics technique for increasing accuracy in severe categorization of road accidents

Publisher: IEEE

Cite This PDF

K. Mohana Prasad ; R. Sabitha ; K. Muthukumar All Authors

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286558

▼ ISBN Information:

Electronic ISBN:978-1-5090-3800-8

Print on Demand(PoD) ISBN:978-1-5090-3801-5

Publisher: IEEE

Conference Location: Chennai, India

Predictive maintenance for wind turbine diagnostics using vibration signal analysis based on collaborative recommendation approach

Publisher: IEEE

Cite This

PDF

Gopi Krishna Durbhaka ; Barani Selvaraj All Authors

▼ ISBN Information:

Electronic ISBN:978-1-5090-2029-4

USB ISBN:978-1-5090-2028-7

Print on Demand(PoD) ISBN:978-1-5090-2030-0

Publisher: IEEE

Conference Location: Jaipur, India

Power factor improvement of compact fluorescent lightings with valley-fill circuit

Publisher: IEEE

Cite This

PDF

S. M. Shyni ; M. W. Abitha ; C. Bhuvaneswari **All Authors**

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290436

▼ ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 International Conference...

Power factor correction in fuzzy based brushless DC motor fed by bridgeless buck boost converter

Publisher: IEEE
Cite This PDF

M. Kavitha ; V. Sivachidambaranathan All Authors

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290426

▼ ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 IEEE International Confe...

Optimized test data generation over suspicious implementation of oracle problem

Publisher: IEEE

Cite This PDF

J Albert Mayan ; M D Anto Praveena ; Telkar Bharath Rao ; Uravakonda Uday Sagar All Authors

Published in: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)

Date of Conference: 21-22 September 2017

Date Added to IEEE *Xplore*: 21 June 2018

INSPEC Accession Number: 17860829

DOI: 10.1109/ICPCSI.2017.8392179

Publisher: IEEE

Conference Location: Chennai, India

▼ **ISBN Information:**

Electronic ISBN:978-1-5386-0814-2

Print ISBN:978-1-5386-0813-5

CD:978-1-5386-0812-8

Print on Demand(PoD) ISBN:978-1-5386-0815-9

Abstract

Microstrip Antenna design has experienced more development in the past period of years and still is subjected to more development. There are different kind of microstrip antenna which can be used for many handheld devices and in the communication devices like satellite link, radar system, radio and cellular mobiles. In this paper the behavior of microstrip antenna is analyzed with two types of electromagnetic band-gap structure such as frequency selective Structure (FSS), and photonic band-gap (PBG) structure. The major characteristics of EBG structure are to show the band gap features in the suppression of surface-wave propagation. This aspect helps to give better antenna performance such as increasing the gain and reducing back radiation. In particular, the distance between EBG cells and the patch is free from the outside control of the cell period, which can be arbitrarily selected, and the final setup offers footprint reduction. With these factors the Gain, bandwidth, and the Voltage standing wave ratio (VSWR) are analyzed using the FEM based EM simulator HFSS.

Conferences > 2017 International Conference...

Novel non-isolated high step-down three phase interleaved DC-DC converter with capacitor voltage division techniques

Publisher: IEEE

Cite This PDF

A. Ramesh Babu ; Raghavendiran ; Barnabas Paul Glady ; G. T. Sundar Rajan All Authors

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICCPCT.2017.8074399

▼ ISBN Information:

Electronic ISBN:978-1-5090-4967-7

Print on Demand(PoD) ISBN:978-1-5090-4968-4

Publisher: IEEE

Conference Location: Kollam, India

Conferences > 2017 International Conference...

Neoteric AT for the visually challenged

Publisher: IEEE

Cite This PDF

P Sivagami ; M Pushpavalli ; N Peter Shyamal All Authors

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290360

▼ ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Advances in Water Resource and Environment Research

June 2017

Edition: 2017 · Publisher: Centre for Remote Sensing and Geoinformatics (CRSG), Sathyabama University · Editor: Kishan Singh Rawat · ISBN: 978-93-83409-39-6

Authors:



Kishan Singh Rawat
Graphic Era University



Anoop Mishra
India Meteorological Department

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/318337757

Advances in Water Resource and Environment Research

Book · June 2017

CITATIONS
0

READS
2,249

2 authors:



Kishan Singh Rawat
Graphic Era University
134 PUBLICATIONS **1,312** CITATIONS
SEE PROFILE



Anoop Mishra
India Meteorological Department
88 PUBLICATIONS **1,049** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Developing Near Real time high resolution precipitation product using multi-satellite sensors for near real time flash flood monitoring View project



Detecting convective clouds over India using multi-spectral satellite observations and quantifying their variability in changing climate View project



Nanomedicine as a Newly Emerging Approach Against Multidrug-Resistant Tuberculosis (MDR-TB)

Muthupandian Saravanan (Mekelle University, Ethiopia & Saveetha Dental College, Saveetha Institute of Medical and Technical Sciences (SIMATS), India), Kebret Duche (Mekelle University, Ethiopia), Tsehaye Asmelash (Mekelle University, Ethiopia), Araya Gebreyesus (Mekelle University, Ethiopia), Anima Nanda (Sathyabama University, India), and Selvaraj Arokiyaraj (Seoul National University, South Korea)

Source Title: Biomedical Engineering: Concepts, Methodologies, Tools, and Applications
Copyright: © 2018 | Pages: 20
DOI: 10.4018/978-1-5225-3158-6.ch041

Conferences > 2017 International Conference...

Multi carrier PWM switching technique for multilevel inverter for various loads

Publisher: IEEE

Cite This PDF

V. Geetha ; V. Sivachidambaranathan All Authors

▼ ISBN Information:
Electronic ISBN:978-1-5090-4324-8
Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 International Conference...

Monitoring Fishy activity of the user in social networking

Publisher: IEEE

Cite This PDF

Mercy Paul Selvan ; Renuka Selvaraj **All Authors**

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070737

▼ ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 Third International Conf...

Mining high utilities patterns in a medical dataset

Publisher: IEEE

Cite This PDF

R. Aroul Canessane ; S. Vaishnavi ; G. S. Vishalini All Authors

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261273

▼ ISBN Information:

Electronic ISBN:978-1-5090-4855-7

Print on Demand(PoD) ISBN:978-1-5090-4856-4

Publisher: IEEE

Conference Location: Chennai, India

Light weight cryptographic algorithms for medical internet of things (IoT) - a review

Publisher: IEEE

Cite This

PDF

D. Jamuna Rani ; S. Emalda Roslin **All Authors**

▼ ISBN Information:
Electronic ISBN:978-1-5090-4556-3
Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Conferences > 2017 IEEE International Confe...

IOT clinic-Internet based patient monitoring and diagnosis system

Publisher: IEEE

Cite This PDF

Sindu Divakaran ; Lavanya Manukonda ; N Sravya ; M Melinda Morais ; P Janani All Authors

Published in: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)

Date of Conference: 21-22 September 2017

Date Added to IEEE *Xplore*: 21 June 2018

INSPEC Accession Number: 17860987

DOI: 10.1109/ICPCSI.2017.8392243

Publisher: IEEE

Conference Location: Chennai, India

▼ **ISBN Information:**

Electronic ISBN:978-1-5386-0814-2

Print ISBN:978-1-5386-0813-5

CD:978-1-5386-0812-8

Print on Demand(PoD) ISBN:978-1-5386-0815-9

Conferences > 2017 IEEE International Confe...

Intelligent walker with obstacle detection technology for visually challenged people

Publisher: IEEE
Cite This PDF

S. Krishnakumar ; Banani Mridha ; M. J. Nancy Naves ; K. Kowsalya All Authors

Published in: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)

Date of Conference: 21-22 September 2017

Date Added to IEEE *Xplore*: 21 June 2018

INSPEC Accession Number: 17860841

DOI: 10.1109/ICPCSI.2017.8392244

Publisher: IEEE

Conference Location: Chennai, India

▼ **ISBN Information:**

Electronic ISBN:978-1-5386-0814-2

Print ISBN:978-1-5386-0813-5

CD:978-1-5386-0812-8

Print on Demand(PoD) ISBN:978-1-5386-0815-9

Conferences > 2017 International conference...

Intelligent power saving system using PIR sensors

Publisher: IEEE

Cite This PDF

R Harikrishnan ; P Sivagami All Authors

Date Added to IEEE *Xplore*: 18 December 2017

DOI: 10.1109/ICECA.2017.8212729

Publisher: IEEE

Conference Location: Coimbatore, India

▼ ISBN Information:

Electronic ISBN:978-1-5090-5686-6

Print ISBN:978-1-5090-5685-9

DVD ISBN:978-1-5090-5684-2

Print on Demand(PoD) ISBN:978-1-5090-5687-3

Conferences > 2017 Third International Conf...

Improving the efficiency for extracting high utility sequential pattern

Publisher: IEEE

Cite This PDF

Raushan Kumar ; R. Aroul Canessane All Authors

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261417

▸ ISBN Information:

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Implementation of template independent web news extraction approach, noise removal and structured data detection to improve search for location based services

Publisher: IEEE

Cite This PDF

R. Jeberson Retna Raj ; P. SundeepTeja ; B. Suryanarayan ; T. Sasipraba All Authors

Date Added to IEEE *Xplore*: 26 October 2017

DOI: 10.1109/ICPEDC.2017.8081074

Publisher: IEEE

Conference Location: Chennai, India

▼ ISBN Information:

Electronic ISBN:978-1-5090-4679-9

DVD ISBN:978-1-5090-4677-5

Print on Demand(PoD) ISBN:978-1-5090-4680-5

Conferences > 2016 International Conference...

Implementation of Imperialist Competitive Algorithm for optimal allocation of FACTS devices to enhance the power system performance

Publisher: IEEE

Cite This PDF

R. Vanitha ; J. Baskaran ; S. Kamalsakthi All Authors

▼ ISBN Information:

Electronic ISBN:978-1-5090-0901-5

CD:978-1-5090-0900-8

Print on Demand(PoD) ISBN:978-1-5090-0902-2

Publisher: IEEE

Conference Location: Melmaruvathur, India

Implementation of high voltage gain interleaved boost converter combined with Switched Capacitor for the low power application

Publisher: IEEE

Cite This

PDF

V.Senthil Nayagam ; L. Premalatha **All Authors**

▼ ISBN Information:
Electronic ISBN:978-1-5090-0901-5
CD:978-1-5090-0900-8
Print on Demand(PoD) ISBN:978-1-5090-0902-2

Publisher: IEEE

Conference Location: Melmaruvathur, India

Implementation of coplanar approximate adders in QCA

Publisher: IEEE Cite This PDF

B. Aravinth ; L. Jegan Antony Marcilin All Authors

▼ ISBN Information:
Electronic ISBN:978-1-4673-9338-6
USB ISBN:978-1-4673-9337-9
Print on Demand(PoD) ISBN:978-1-4673-9339-3

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Image based password authentication system for banks

Publisher: IEEE

Cite This PDF

S. Sukanya ; M. Saravanan All Authors

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070764

▼ ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 Third International Conf...

Home and office automation system-survey

Publisher: IEEE

Cite This PDF

V. Gobinath ; L. Mary Gladence ; S. Hari Haran ; R. Sathiyavathi ; V. Maria Anu All Authors

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261428

▼ ISBN Information:

Electronic ISBN:978-1-5090-4855-7

Print on Demand(PoD) ISBN:978-1-5090-4856-4

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

High utility sequential pattern mining using intelligent technique

Publisher: IEEE

Cite This

PDF

Daison Joseph ; Gaurav Kumar Bansal ; P. Asha **All Authors**

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICCPCT.2017.8074310

▼ ISBN Information:

Electronic ISBN:978-1-5090-4967-7

Print on Demand(PoD) ISBN:978-1-5090-4968-4

Publisher: IEEE

Conference Location: Kollam, India

Conferences > 2017 International Conference...

Graphical passwords implies on tolerance password, image choice, and puzzle login security

Publisher: IEEE

Cite This PDF

L. Gunaseeli ; R. Aroul Canessane All Authors

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070758

▼ ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Fuzzy-based cluster head election mechanisms for wireless sensor networks: A review

Publisher: IEEE

Cite This PDF

S Ramakrishnan ; S. Prayala Shyry All Authors

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070714

▼ ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Fault behaviour pattern analysis and recognition

Publisher: IEEE Cite This PDF

Gopi Krishna Durbhaka ; S. Barani **All Authors**

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-1987-8

Print on Demand(PoD) ISBN:978-1-5090-1988-5

Publisher: IEEE

Conference Location: Kochi, India

Conferences > 2017 IEEE International Confe...

Entry and exit monitoring using license plate recognition

Publisher: IEEE

Cite This PDF

A. Yovan Felix ; A. Jesudoss ; J. Albert Mayan All Authors

Date Added to IEEE *Xplore*: 30 October 2017

DOI: 10.1109/ICSTM.2017.8089156

Publisher: IEEE

Conference Location: Chennai, India

▼ ISBN Information:

Electronic ISBN:978-1-5090-5905-8

Print ISBN:978-1-5090-5918-8

CD:978-1-5090-5904-1

Print on Demand(PoD) ISBN:978-1-5090-5906-5

Conferences > 2017 2nd International Confer...

Enhancement of distinct dynamic broadcasting technique in architecture of object detection and tracking with cloud

Publisher: IEEE
Cite This PDF

T. Mahalingam ; P. Jeyanthi All Authors

Date Added to IEEE *Xplore*: 11 July 2017

DOI: 10.1109/ICCCT2.2017.7972248

▼ ISBN Information:

Electronic ISBN:978-1-5090-6221-8

CD:978-1-5090-6220-1

Print on Demand(PoD) ISBN:978-1-5090-6222-5

Publisher: IEEE

Conference Location: Chennai, India

Enhanced home automation system using Internet of Things

Publisher: IEEE

Cite This PDF

S.L.S. Sri Harsha ; S. Chakrapani Reddy ; S. Prince Mary All Authors

16 Paper Citations
1266 Full Text Views

Abstract

Document Sections

I. Introduction

II. System Design

III. Hardware Design

IV. Software Design

V. Conclusion and Future Enhancements

Authors

Abstract:

The more the technology advances into the future, the more it makes the life of people depending on it easier and one of such technologies is Automation. The term Automation can be coupled with-no effort, ease of performing a task and less human involvement. Internet of Things is such other term which envisions every object around us as an integral part of the Internet. In this paper, we suggest a highly intractable and environmentally sustainable form of Home Automation System using Internet of Things a means to control the appliances at home via a device with access to the Internet. The key components of this system are a pocket sized microprocessor-Raspberry Pi and a microcontroller-Arduino Uno and an Android application to visualize the data provided by the Raspberry Pi and also to send, receive and process the requests. The Raspberry Pi acts as the brain of this system, processing the requests, responding to the requests made by the Android application, communicating with the Arduino and also acts as a server to store the data given by the sensors. All the sensors and actuators are connected to the Arduino which is connected to the Pi using a USB cable. Our main objective of developing this model is to create a home automation system which interacts with the user through various push notifications based on concerned parameters which is also eco-friendly.

Date Added to IEEE *Xplore*: 05 October 2017

DOI: 10.1109/I-SMAC.2017.8058302

▼ ISBN Information:

Electronic ISBN:978-1-5090-3243-3

Print ISBN:978-1-5090-3242-6

DVD ISBN:978-1-5090-3241-9

Print on Demand(PoD) ISBN:978-1-5090-3244-0

Publisher: IEEE

Conference Location: Palladam, India

Abstract

WSN comprises of sensor nodes distributed spatially to accumulate and transmit measurements from environment through radio communication. It utilizes energy for all its functionality (sensing, processing, and transmission) but energy utilization in case of transmission is more. Data compression can work effectively for reducing the amount of data to be transmitted to the sink in WSN. The proposed compression algorithm i.e. Energy Efficient Deflate (EEDeflate) along with fuzzy logic works effectively to prolong the life of Wireless Sensor Network. EEDeflate algorithm saves 7% to 10% of energy in comparison with the data transmitted without compression. It also achieves better compression ratio of average 22% more than Huffman and 8% more than Deflate compression algorithm. This improvements in terms of compression efficiency allows saving energy and therefore extends the life of the sensor network.

Energy Efficient Deflate (EEDeflate) Compression for Energy Conservation in Wireless Sensor Network

Pramod Ganjewar ✉, S. Barani & Sanjeev J. Wagh

Electronic Circuit Breaker for overload protection

Publisher: IEEE Cite This PDF

P. Abirami ; Merin Lizbeth George **All Authors**

9 Paper Citations | 1 Patent Citation | 634 Full Text Views

Abstract

Document Sections

I. Introduction

II. Block Diagram

III. Simulation Work

IV. Hardware Description

Abstract:

This paper proposes an Electronic Circuit Breaker (ECB), trips the power supply instantaneously in overload condition Mechanically operated circuit breakers like MCB is based on very slow action thermal bimetal lever trip mechanism. The activation time of MCB rely on the overloaded percentage of the fault. The concept of electronic circuit breaker is introduced to know that the thermally operated circuit breakers such as MCBs take longer time to trip. For sensitive loads it is necessary to activate the circuit breaker within the précised time. The current flow through the series element is sensed by ECB and the corresponding voltage drop and the preset voltage is compared with a level comparator and the resultant signals are given to the microcontroller through a MOSFET which drives the relay to trip the load. Electronic circuit breaker is operated instantaneously with the tripping of relay.

Date Added to IEEE *Xplore*: 01 September 2016

DOI: 10.1109/ICCPEIC.2016.7557324

▼ ISBN Information:

Electronic ISBN:978-1-5090-0901-5

CD:978-1-5090-0900-8

Print on Demand(PoD) ISBN:978-1-5090-0902-2

Publisher: IEEE

Conference Location: Melmaruvathur, India

Conferences > 2017 IEEE International Confe...

Early diagnosis of prostate cancer using image processing techniques

Publisher: IEEE

Cite This PDF

J Bethanney Janney ; J Jessie Christilda ; S Sindhuja Mary ; D Haritha **All Authors**

Date of Conference: 21-22 September 2017

Date Added to IEEE *Xplore*: 21 June 2018

ISBN Information:

Electronic ISBN:978-1-5386-0814-2

Print ISBN:978-1-5386-0813-5

CD:978-1-5386-0812-8

Print on Demand(PoD) ISBN:978-1-5386-0815-9

INSPEC Accession Number: 17860720

DOI: 10.1109/ICPCSI.2017.8392245

Publisher: IEEE

Conference Location: Chennai, India

Abstract

Electric vehicles (EV) and hybrid electric vehicles (HEV) are operated by battery power, which has to be optimized. HEV is the combination of an electric machine and an internal combustion engine (ICE) which is a promising means of reducing emissions and fuel consumption without compromising vehicle functionality and driving performance. The hybrid electric vehicle can be classified according to the way in which power is supplied to drive the train as parallel hybrid electric vehicle (PHEV), series hybrid electric vehicle (SHEV), dual-mode hybrid electric vehicle (DMHEV). In this proposed work, dual-mode hybrid electric vehicle (DMHEV) is used. One of the motors in the drive is a permanent magnet synchronous motor (PMSM). This PMSM serves two operations: As a parallel motor with an existing motor and the winding of PMSM acts as a boost inductor. A dual converter is proposed which acts as a DC-to-AC converter when PMSM is working as a motor. The same converter acts as a boost converter which uses the PMSM inductance in the boost mode. In this work, a total automobile HEV system is simulated using MATLAB. This system has an electrical subsystem which contains a generator, a battery and motors. The dual converter and PMSM are included in the electrical subsystem to improve the performance of the automobile system.



Proceedings of the International Conference on Soft Computing Systems pp 237–249 | Cite as

Dual Converter Multimotor Drive for Hybrid Permanent Magnet Synchronous in Hybrid Electric Vehicle

G. Nagarajan ✉, C. N. Ravi, K. Vasanth, D. Godwin Immanuel & S. D. Sundarsingh Jebaseelan

Conferences > 2017 International Conference...

Double frequency SPWM inverter topology for PV systems

Publisher: IEEE

Cite This PDF

V. Sivachidambaranathan ; V. Geetha **All Authors**

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290476

ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Development of efficient quality preserving invisible watermarking technique to embed both images and data in an image

Publisher: IEEE

Cite This

PDF

Malli B ; Lagishetty Mounica ; Nandhitha. N.M ; Balamurugan. V **All Authors**

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-4556-3

Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Design of low power radix-2 and radix-4 ACSU in Viterbi decoder using HSPICE

Publisher: IEEE

Cite This

PDF

Prakash. S Harsha ; V. Balamurugan **All Authors**

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-4556-3

Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Design of less time delay multiplier using vedic mathematics

Publisher: IEEE

Cite This

PDF

Megha Dogra ; V. Balamurugan **All Authors**

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-4556-3

Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Design of highly compressed NB PHY transceiver of IEEE 802.15.6 based on adaptive modulation scheme

Publisher: IEEE Cite This PDF

M. Mercy Twinkle ; G. Jegan All Authors

ISBN Information:

Electronic ISBN:978-1-5090-4556-3

Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India

Conferences > 2017 Third International Conf...

Design of contactless power transfer for artificial heart

Publisher: IEEE

Cite This

PDF

Lidhiya Daniel ; Thukkaram Sudhakar ; Sheeba Abraham **All Authors**

Date Added to IEEE *Xplore*: 26 October 2017

DOI: 10.1109/ICBSII.2017.8082287

ISBN Information:

Electronic ISBN:978-1-5090-4980-6

Print on Demand(PoD) ISBN:978-1-5090-4981-3

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Design and analysis of logic gates using single electron nano-devices

Publisher: IEEE

Cite This PDF

S. Rajasekaran ; G. Sundari **All Authors**

Date of Conference: 23-23 September 2017

Date Added to IEEE *Xplore*: 19 April 2018

ISBN Information:

Electronic ISBN:978-1-5386-2138-7

DVD ISBN:978-1-5386-2137-0

Print on Demand(PoD) ISBN:978-1-5386-2139-4

INSPEC Accession Number: 17719997

DOI: 10.1109/ICAETGT.2017.8341459

Publisher: IEEE

Conference Location: Coimbatore, India

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/321343379

Contingency Analysis Reliability Evaluation of Small-Signal Stability Analysis

Chapter *in* Lecture Notes in Electrical Engineering · January 2018

DOI: 10.1007/978-981-10-4394-9_24

CITATIONS
0

READS
85

3 authors, including:



Murugan Arumugham
Sathyabama Institute of Science and Technology
8 PUBLICATIONS 23 CITATIONS

SEE PROFILE



Jayaprakash Subramanian
Sathyabama Institute of Science and Technology
25 PUBLICATIONS 93 CITATIONS

SEE PROFILE

Conferences > 2017 Third International Conf...

Condition monitoring and sensor based robust protection for inverter fed induction motor drive system

Publisher: IEEE

Cite This PDF

C. Sravani ; V. L. K. Chaitanya ; M. W. Abitha **All Authors**

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261403

ISBN Information:

Electronic ISBN:978-1-5090-4855-7

Print on Demand(PoD) ISBN:978-1-5090-4856-4

Publisher: IEEE

Conference Location: Chennai, India

Comparative analysis of scheduling and load balancing algorithms in cloud environment

Publisher: IEEE Cite This PDF

K. Pradeep ; T. Prem Jacob All Authors

▼ **ISBN Information:**

Electronic ISBN:978-1-5090-5240-0

DVD ISBN:978-1-5090-5239-4

Print on Demand(PoD) ISBN:978-1-5090-5241-7

Publisher: IEEE

Conference Location: Kumaracoil, India

2581 Accesses | 2 Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 394)

Abstract

This paper implements the regenerative braking concept with a bidirectional DC–DC converter-fed DC motor to charge the ultracapacitor and also to reduce power consumption. This paper uses the ultracapacitor power source for fast current changes and while battery responds for slow current changes which in turn increases the lifetime of the battery. Energy developed at the time of braking is not useless; it is reutilized to store the energy in the ultracapacitor. An experimental prototype is developed. The advantage is that state of charge (SOC) of the battery and ultracapacitor voltage and current and voltage are within the limits.

Conferences > 2017 International Conference...

Authenticated anonymous secure on demand routing protocol in VANET (Vehicular adhoc network)

Publisher: IEEE

Cite This PDF

K. Logeshwari ; L. Lakshmanan **All Authors**

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070730

ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

Conferences > 2017 International Conference...

Attendance tracking using Wi-Fi

Publisher: IEEE

Cite This

PDF

Basu Kumar Swamy ; R Vanitha ; Deepak Kumar **All Authors**

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290407

ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

New Trends and Issues Proceedings on Humanities and Social Sciences

Volume 4, Issue 11, (2017) 93-106

ISSN : 2547-881
www.prosoc.eu

Selected Paper of 6th World Conference on Design and Arts (WCDA 2017), 29 June – 01 July 2017, University of Zagreb, Zagreb – Croatia

Art of Facilitating 'Problem-Driven Outcomes' in an Architectural Design Studio

Arulmalar Ramaraj [a*], Sathyabama University, Chennai 600119, India.
Jothilakshmy Nagammal [b], Sathyabama University, Chennai 600119, India.

Suggested Citation:
Ramaraj, A. & Nagammal, J. (2017). Art of facilitating 'Problem driven outcomes' in an architectural design Studio. *New Trends and Issues Proceedings on Humanities and Social Sciences.* [Online]. 4(11), 93-106. Available from: www.prosoc.eu

Selection and peer review under responsibility of Prof. Dr. Ayse Cakır Ilhan, Ankara University, Turkey.
©2017 SciencePark Research, Organization & Counseling. All rights reserved.

Application of computational chemistry for adsorption studies on metal-organic frameworks used for carbon capture

A. Annam Renita and V. Sivasubramanian

A. Annam Renita

Department of Chemical Engineering, Sathyabama Institute of Science and Technology, Chennai600119, India

Email author

Search for this author in:

De Gruyter | Google Scholar

is content. Please log in with your institutional or personal account if you should have
a limited preview of this publication:

Computational chemistry is invaluable in calculating macroscopic and microscopic details of systems application in chemical industries which are involved in carbon capture through precombustion, post-combustion and oxy combustion technologies. This review discusses the role of computational chemistry for adsorption studies of metal–organic frameworks (MOFs) which can be utilized for carbon capture. Principles of quantum mechanics–molecular mechanics are used to devise the electrostatic charges and isotherm parameters on the MOFs. MOFs for carbon capture which can be compatible and which can withstand the severity in chemical industries can be effectively studied using grand canonical Monte Carlo simulation by selecting appropriate force fields.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 6, Special Issue 3, November 2017

ANFIS Based STATCOM for Compensation of Reactive Power and Harmonics with Cascaded H-Bridge Multilevelinverter

K.A.Latha, Dr. S.D.Sundarsingh Jebaseelan

PG Scholar, Department of EEE, Sathyabama University, Chennai, India

Associate Professor, Department of EEE, Sathyabama University, Chennai, India

ABSTRACT: This paper investigates the improvement of voltage stability using STATCOM based on CHB Multilevel Inverter. In a transmission line if there is a sudden load change occurs it lead to power quality issues . STATCOM is connected in parallel with the transmission line and eliminates the voltage distortion. Normal inverter produces the non-sinusoidal waveform so the power quality may get reduces so the multilevel inverter is implemented in STATCOM to overcome this issue. Additionally, DC link voltage stability is improved using the ANFIS controller.

KEYWORDS: FACTS, STATCOM, Adaptive fuzzy Inference system **(ANFIS),** cascaded H bridge multilevel inverter **(CHB),** power quality **(PQ)** Total harmonic distortion **(THD)**

I. INTRODUCTION

Power distribution system should provide an un-interrupted supply to their consumers with smooth sinusoidal voltage and constant frequency Distribution systems consist of several non- linear loads that considerably affect the quality of power. power quality can also affected by the events like capacitor switching ,motor starting

PQ problem is defined as any established problem in voltage, current or leading to frequency deviations. It will result in failure or mal–operation of the customer equipment

An novel approach for segmentation using brain images

Publisher: IEEE

Cite This

PDF

V. Ulagamuthalvi ; G. Kulanthaivel All Authors

Date Added to IEEE *Xplore*: 24 July 2017

DOI: 10.1109/ICCICCT.2016.7988006

ISBN Information:

Electronic ISBN:978-1-5090-5240-0

DVD ISBN:978-1-5090-5239-4

Print on Demand(PoD) ISBN:978-1-5090-5241-7

Publisher: IEEE

Conference Location: Kumaracoil, India

An improvised X-filling PRL strategy for test data compression

Publisher: IEEE

Cite This

PDF

Poorani S. ; Muthiah M.A. **All Authors**

134
Full
Text Views

Abstract

Document Sections

I. Introduction

II. Proposed Technique

III. Experimental Results and Comparisons

IV. Conclusion

Abstract:

A huge voluminous of test data creates a major problem in the testing of system-on-a-chip (SOC) design. The advent of compressing test data is mainly due to an inadequate channel capacity and the restricted memory of Automatic Test Equipment (ATE) to store the test data. To overcome these problems, an efficient approach is presented here, namely Cyclic shift based pattern run length (PRL) coding for reducing the volume of test vectors. This new type of Pattern Run Length coding works by compressing the consecutive patterns based on the possible occurrences of four different cases when partitioning the test vector into number of fixed patterns. The experimental results for ISCAS'89 benchmarks show that this proposed technique results in greater compression ratio than former techniques. In addition to, the analytical results show that the cyclic shift based PRL is more suitable when the segment length k increases.

▼ ISBN Information:
Electronic ISBN:978-1-5090-4556-3
Print on Demand(PoD) ISBN:978-1-5090-4557-0

Publisher: IEEE

Conference Location: Coimbatore, India



Artificial Intelligence and Evolutionary Computations in Engineering Systems pp 1195–1203 | Cite as

Bidirectional Buck–Boost Converter-Fed DC Drive

P. Karthikeyan & V. Siva Chidambaranathan

Conference paper | First Online: 06 February 2016

2577 Accesses | **2** Citations

Part of the Advances in Intelligent Systems and Computing book series (AISC,volume 394)

Abstract

This paper implements the regenerative braking concept with a bidirectional DC–DC converter-fed DC motor to charge the ultracapacitor and also to reduce power consumption. This paper uses the ultracapacitor power source for fast current changes and while battery responds for slow current changes which in turn increases the lifetime of the battery. Energy developed at the time of braking is not useless; it is reutilized to store the energy in the ultracapacitor. An experimental prototype is developed. The advantage is that state of charge (SOC) of the battery and ultracapacitor voltage and current and voltage are within the limits.

Access via your institution →

Chapter EUR 29.95
Price includes VAT (India)

- DOI: 10.1007/978-81-322-2656-7_109
- Chapter length: 9 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

eBook EUR 234.33

Softcover Book EUR 279.99

Learn about institutional subscriptions

Alleviate the voltage gain of high step-up DC to DC converter using quasi active switched inductor structure for renewable energy

Publisher: IEEE

Cite This PDF

V. Balasubramanian ; V. Senthil Nayagam ; Jayarama Pradeep **All Authors**

1 Paper Citation

293 Full Text Views

Abstract

Document Sections

I. Introduction

II. Principle of Operation

III. Voltage Gain Expression

IV. Simulation Block Diagram of the Modified High Step-Up DC to DC Converter QASL Network with Pi Filter

V. Simulation Output and Comparisons Between Conventional and Modified

Abstract:

The demand for non-isolated high step-up DC-DC-dc converters in applications such as DC backup energy systems for UPS and photovoltaic has been gradually increasing. Generally, Boost converter should operate at high duty cycle in order to get the high-output voltage but it generates more EMI, conduction loss and increase the switch rating. High voltage gain DC-DC converter implemented quasi-active switched-inductor structure for a non conventional system is included in this paper. The PV system is available in the proposed system as well as two coupled inductors with magnetic core, two capacitors, two active switches, three diodes and PI filter [1]. The input supply is connected in shunt with primary part of the coupled inductor for charging. Meanwhile the charged coupled inductor is discharged in series manner along with the source and the capacitor in order to get efficient high output voltage with proper duty cycle is selected. The diode capacitor set will act as a clamp circuit so that the voltage conversion ratio by the coupled inductor structure is further improved to high level with less stress on the switch. In order to reduce the conduction loss by the switches used we preferred MOSFET switches with very low on state resistance. The reverse recovery problem is overcome with the help of active clamp circuit [2]-[3]. The proposed circuit model is simulated using MATLAB.

Published in: 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC)

Date Added to IEEE *Xplore*: 15 February 2018

DOI: 10.1109/ICCPEIC.2017.8290482

▼ ISBN Information:

Electronic ISBN:978-1-5090-4324-8

Print on Demand(PoD) ISBN:978-1-5090-4325-5

Publisher: IEEE

Conference Location: Melmaruvathur, India

Advanced automation control in an ambulance under emergency condition

Publisher: IEEE Cite This PDF

Leila Sai Krishna ; Samineni Vijay Chowdary ; M. Pushpavalli ; P. Sivagami All Authors

2 Paper Citations
183 Full Text Views

Abstract
Document Sections
I. Introduction
II. Related Work
III. Block Diagram of Proposed System
IV. Simulation Results
V. Hardware Results
Show Full Outline

Abstract:
Currently, India is facing huge traffic congestion and the traffic disturbance in many major cities around is very severe. Mainly in urban areas, most of the people are using cars as transport when they go out. Due to this traffic blockage, there is a rise in road accidents which direction to a ruin of individual human lives. To avoid this we have implemented a scheme which can control the traffic signals automatically in its path way and reduce the amount of traffic congestion at the signal. The ambulance is implemented with embedded system units, which finds the accident spot and delivers the spot to the close by ambulance through GPS. The ambulance guides the traffic lights in the roadway to the hospital and it also checks the next consecutive signal to shorten the time loss. The vehicle unit is equipped with vibration sensor to determine the vibration if it exceeds the level. The embedded unit sense the accident happens and its sends the neighborhood ambulance unit through wireless transmission. RFID reader is needed to extricate the ambulance from other vehicles.

Published in: 2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)

Date Added to IEEE *Xplore*: 18 January 2018

DOI: 10.1109/ICONSTEM.2017.8261399

▼ ISBN Information:

Electronic ISBN:978-1-5090-4855-7

Print on Demand(PoD) ISBN:978-1-5090-4856-4

Publisher: IEEE

Conference Location: Chennai, India

A survey paper on big data analytics

Publisher: IEEE Cite This PDF

M. D. Anto Praveena ; B. Bharathi **All Authors**

25 Paper Citations

2864 Full Text Views

Abstract

Document Sections

I. Introduction

II. Big Data-An Overview

III. Big Data Technologies

IV. Big Data Challenges

V. Big Data Analysis Algorithms

Show Full Outline

Authors

Abstract:
In recent years, the internet application and communication have seen a lot of development and reputation in the field of Information Technology. These internet applications and communication are continually generating the large size, different variety and with some genuine difficult multifaceted structure data called big data. As a consequence, we are now in the era of massive automatic data collection, systematically obtaining many measurements, not knowing which one will be relevant to the phenomenon of interest. For example, E-commerce transactions include activities such as online buying, selling or investing. Thus they generate the data which are high in dimensional and complex in structure. The traditional data storage techniques are not adequate to store and analyses those huge volume of data. Many researchers are doing their research in dimensionality reduction of the big data for effective and better analytics report and data visualization. Hence, the aim of the survey paper is to provide the overview of the big data analytics, issues, challenges and various technologies related with Big Data.

Published in: 2017 International Conference on Information Communication and Embedded Systems (ICICES)

Date Added to IEEE *Xplore*: 19 October 2017

DOI: 10.1109/ICICES.2017.8070723

▼ ISBN Information:

Electronic ISBN:978-1-5090-6135-8

Print on Demand(PoD) ISBN:978-1-5090-6136-5

Publisher: IEEE

Conference Location: Chennai, India

A survey on data redundancy and node charge preserved underwater acoustic sensor network

Publisher: IEEE

Cite This

PDF

R. Raymond ; B. Linston Joe Savio ; R. M. Gomathi ; K. Priya **All Authors**

Department of Information Technology, Sathyabama University, Chennai, Tamil Nadu, India

1 Paper Citation

Text Views

Abstract

Document Sections

I. Introduction

II. Basic Architecture of Underwater Sensor Network

III. Related Work

IV. Conclusion

Abstract:

On this paper, the Topological formation of UWAS (under Water Acoustic Sensor) networks encompass AUV nodes and control sink, whereas control sinks are commonly floating buoys at the control of water inflated by way of pumps. These AUVs are without problems maintained and collect the sensed facts in underwater environments. The more than one AUVs for amassing the sensed records in underwater environments. The Mobicast routing protocol is used for onshore sensor networks and the primary intention of this paper is to get rid of the hole problem by creating zone of Forwarding (ZOF) areas. The proposed tactics use a static sensor node which collects the accrued statistics from different nodes in its place and send those accumulated facts to AUV whilst it enters 3D ZOR&ZOF thereby decreasing the strength used for transmission.

Published in: 2017 International Conference on Communication and Signal Processing (ICCSP)

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286591

ISBN Information:

Electronic ISBN:978-1-5090-3800-8

Print on Demand(PoD) ISBN:978-1-5090-3801-5

Publisher: IEEE

Conference Location: Chennai, India

Energy Procedia
Volume 117, June 2017, Pages 991-998



A Novel bridgeless SEPIC Converter for Power Factor Correction

R. MEENADEVI [a], Dr. L. PREMALATHA [b]

[a] Research scholar, Sathyabama University, Chennai, Tamil Nadu, India
[b] Professor/ School of Electrical Engineering, Vellore Institute of Technology (VIT), Chennai, Tamil Nadu, India

Available online 14 September 2017, Version of Record 14 September 2017.

Check for updates

Show less

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.egypro.2017.05.220

Get rights and content

Under a Creative Commons license

Open access

A novel approach for time series data forecasting based on ARIMA model for marine fishes

Publisher: IEEE

Cite This PDF

M O D Rizwan ; R. Jeberson Retna Raj ; M. Vasudev All Authors

2 Paper Citations
209 Full Text Views

Abstract

Document Sections

I. Introduction

II. Literature Survey

III. Proposed Methodology

IV. Algorithm for Predictive Processing

V. Result and Analysis

Show Full Outline

Abstract:
Forecasting of the time series data is a challenge as its details are more complex in nature. Climate change is a global issue as it influences the environment and it impacts directly affecting the human and the marine species. Global warming is a big threat and it reflects in sea temperature. Due to the rising sea temperature, the fishes like sardine and pelagic are not getting the living sea environment and the annual catchment of these fishes are reduced. The fishes are drifted from one place to other. This paper focuses on forecasting these fishes based on the catchment of historical data. The proposed approach incorporates the Multivariate Imputation by Chained Equations (MICE) is used to determine the missing values. The Auto Regressive Integrated Moving Average (ARIMA) model is blended for predicting the upfront values. The proposed approach is tested with various parameters and the test results shows its efficiency.

Published in: 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)

Date Added to IEEE *Xplore*: 14 December 2017

DOI: 10.1109/ICAMMAET.2017.8186707

▼ ISBN Information:

Electronic ISBN:978-1-5090-3378-2

Print on Demand(PoD) ISBN:978-1-5090-3379-9

Publisher: IEEE

Conference Location: Chennai, India

A novel approach for securing data de-duplication methodology in hybrid cloud storage

Publisher: IEEE

Cite This

PDF

Kameswari Bhaskar ; R Jayashree ; R. Sathiyavathi ; L. Mary Gladence ; V. Maria Anu **All Authors**

3 Paper Citations

102 Full Text Views

Abstract

Document Sections

I. Introduction

II. Related Work

III. Proposed System

IV. System Architecture

V. Result and Discussion

Authors

Figures

Abstract:

Secured de-duplication, with the approach of distributed computing, has pulled in much consideration as of late from research group. The most imperative and well known cloud administration is information stockpiling. With a specific end goal to save the security of the information holder, the information is normally put away in the encoded frame. Conventional de-duplication doesn't take a shot at scrambled information. This paper speaks to private information de-duplication, one of the effective pressure strategy, in half and half cloud. This strategy is generally utilized as a part of cloud to limit the information storage room and transmission capacity with a specific end goal to ad lib the effectiveness. The secrecy is the greatest test in de-duplication system. The delicate information must be shielded from outer assaults, hence these information needs to experience encryption before outsourcing. To defeat the security shortcoming, this paper gives the model to a secured de-duplication in a group cloud.

Published in: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)

Date Added to IEEE *Xplore*: 01 February 2018

DOI: 10.1109/ICIIECS.2017.8275888

▼ ISBN Information:

Electronic ISBN:978-1-5090-3294-5

Print on Demand(PoD) ISBN:978-1-5090-3295-2

Publisher: IEEE

Conference Location: Coimbatore, India

Conferences > 2017 International Conference...

A new multimodal brain tumour image fusion for pre-processing of medical image database

Publisher: IEEE

Cite This PDF

T. Denzil Bosco ; A. Lakshmi Narayanan ; A. Veeramuthu All Authors

Date Added to IEEE *Xplore*: 08 February 2018

DOI: 10.1109/ICCSP.2017.8286687

▼ ISBN Information:

Electronic ISBN:978-1-5090-3800-8

Print on Demand(PoD) ISBN:978-1-5090-3801-5

Publisher: IEEE

Conference Location: Chennai, India

Dr. N. M. NANDHITHA, Dr. N. SELVARASU,
Dr. S. EMALDA ROSLIN, Dr. REKHA CHAKRAVARTHI, M. S. SANGEETHA

A MONOGRAPH ON COMPUTER AIDED INTERPRETATION OF MEDICAL IMAGES IN MATLAB



- ISBN13:9781947498495
- Language:English
- SUPC: SDL134745472
- ISBN10:1947498495
- Author:Dr.N.M.Nandhitha, Dr.N.Selvarasu, Dr.S.Emalda Roslin, Dr.Rekha Chakravarthi, M.S.Sangeetha
- Publisher:Notion Press Media Pvt Ltd
- Binding:Paperback
- Publishing Year:2017
- Pages:44

View PDF | Access through your institution | Purchase PDF

Chapter contents | Book contents

Front matter
Copyright
Contributors
Biography
Preface
Chapter 1. Introduction to Seaweed Polysaccharides
Chapter 2. Isolation and Characterization of Alginate from...
Chapter 3. Extraction of Sulfated Polysaccharides (Fucoida...
Chapter 4. Seaweed Polysaccharide Isolation Using Subcri...
Chapter 5. Ultrasonic-Assisted Extraction of Carrageenan
Chapter 6. Vibrational Spectroscopy of Seaweed Polysacch...
Chapter 7. Green Synthesis of Metal Nanoparticles Using ...
Chapter 8. Chemical Modification of Alginate

ELSEVIER

Seaweed Polysaccharides
Isolation, Biological and Biomedical Applications
2017, Pages 101-109



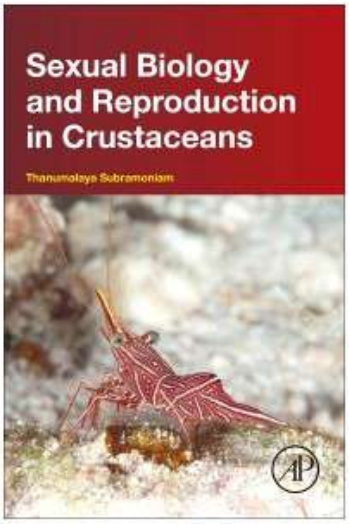
Chapter 7 - Green Synthesis of Metal Nanoparticles Using Seaweed Polysaccharides

Gobi S. Kaliaraj [*], Bharathiraja Subramaniyan [**], Panchanathan Manivasagan [†], Se-Kwon Kim [†]

[*] Sathyabama University, Chennai, Tamil Nadu, India
[**] Annamalai University, Porto Novo, Tamil Nadu, India
[†] Pukyong National University, Busan, South Korea

Citation preview

TY - BOOK
AU - Gandhi, Mercy
PY - 2016/08/20
SP -
SN - ISBN 978-93-83409-19-8
T1 - Delicacy of Digital Literacy
ER -



Preview

Sexual Biology and Reproduction in Crustaceans

1st Edition - September 22, 2016

☆☆☆☆☆ Write a review

Author: Thanumalaya Subramoniam

eBook ISBN: 9780128096062

Paperback ISBN: 9780128093375