

A STUDY ON DATA OUTSOURCING PRACTICES PERFORMED AND IMPLEMENTING THE CRYPTOGRAPHY IN THE INDUSTRY

Submitted in partial fulfilment of the requirements for the award of

MASTER OF BUSINESS ADMINISTRATION

By

JERRY IMMANUEL RAJ P

REGISTER NO:42410020



**DEPARTMENT OF BUSINESS ADMINISTRATION
SCHOOL OF MANAGEMENT STUDIES**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

**Accredited with Grade "A++" by NAAC I 12B Status by UGC I
CATEGORY - 1 UNIVERSITY BY UGC I Approved by AICTE
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI-600119**

December 2023



SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
CATEGORY - 1 UNIVERSITY BY UGC**

SCHOOL OF MANAGEMENT STUDIES

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the Bonafide work of **JERRY IMMANUEL RAJ P 42410020** who carried out the Project Training “**A STUDY ON DATA OUTSOURCING PRACTICES PERFORMED AND IMPLEMENTING THE CRYPTOGRAPHY IN THE INDUSTRY**” under my supervision from October 2023 to December 2023.

Internal guide

External Guide

**Dr. BHUVANESWARI. G, MBA., Ph.D.
Dean–School of Management Studies**

Submitted for Viva voce Examination held on_____

Internal Examiner

External Examiner

DECLARATION

I **JERRY IMMANUEL RAJ.P (42410020)** hereby declare that the Project Report entitled **“A STUDY ON DATA OUTSOURCING PRACTICES PERFORMED AND IMPLEMENTING THE CRYPTOGRAPHY IN THE INDUSTRY”** done by me under the guidance of **DR. GANANATH KHILLA** is submitted in partial fulfilment of the requirements for the award of Master of Business Administration degree.

DATE:

PLACE:

JERRY IMMANUEL RAJ.P

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to Board of Management of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my sincere thanks to **Dr. G. Bhuvaneswari, MBA., Ph.D., Dean - School of Management Studies** and **Dr. A. Palani, M.Com., M.Phil., M.B.A., Ph.D., - Head - School of Management Studies** for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **DR. GANANATH KHILLA** for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **School of Management Studies** who were helpful in many ways for the completion of the Training.

JERRY IMMANUEL RAJ.P

ABSTRACT

Security and privacy in the cloud computing has generated much interest in the research community in recent years. This is due to many advantages at the same time there are available and certain measures to be taken. The confidential documents have been identified as one of the most important asset and it need to be stored and accessed. Usually the encrypted documents are stored on remote cloud servers. There are many existing practices that are proposed to perform security for the data. But still data is being misused or leaked in some instances that provides risky situation to managers of the concern. There need to be proposed a technique that will overcome the risk. This project identifies and analyzes the data outsourcing practices that are followed in the organization and provide them valuable suggestions and recommendations to maintain or improve the data outsourcing practices.

TABLE OF CONTENTS

Chapter No	Title	Page No
	Declaration	i
	Acknowledgement	ii
	Abstract	iii
1	Introduction	1
	1.1 Industry Profile	2
	1.2 Concepts	9
	1.3 Need for the study	12
	1.4 Problem of the study	12
	1.5 Objectives of the study	12
2	Literature Review	13
	2.1 Review of Literature	13

3	Research Methodology	18
	3.1 Research Approach	18
	3.2 Collection of Data	18
	3.3 Limitations	19
4	Data Analysis	20
	4.1 Tables and Charts	20
	4.2 SWOT Analysis	40
5	Summary, Conclusion and Suggestions	41
	5.1. Summary of Findings	41
	5.2 Conclusion	45
	5.3 Suggestions and Recommendations	45
	References	46
	Appendix Questionnaire	

LIST OF TABLES

TABLE NO	TITLE	PAGE NO
4.1.1	Respondents classification based on Age	20
4.1.2	Respondents classification based on Gender	21
4.1.3	Respondents classification based on Educational Qualification	22
4.1.4	Respondents classification based on Experience in Information Security	23
4.1.5	Respondents classification based on Monthly Income (in Rs)	24
4.1.6	Respondents classification based on the algorithm that will be the best suggested	25
4.1.7	Respondents classification based on when the server will be usually upgraded in the company	26
4.1.8	Respondents classification based on starting to upload the confidential data	27
4.1.9	Respondents classification based on ensuring the security of the data in the organization	28
4.1.10	Respondents classification based on expecting the management to provide security of data	29
4.1.11	Respondents classification based on the algorithm that is presently used for secure data outsourcing	30
4.1.12	Respondents classification based on the satisfaction with the performance of the data outsourcing	31
4.1.13	Respondents classification based on the satisfaction with the depth of the working of the security algorithm	32
4.1.14	Respondents classification based on the satisfaction with the training methods of the working of the algorithm	33

4.1.15	Respondents classification based on the satisfaction with the quality of training of the data outsourcing process	34
4.1.16	Respondents classification based on the satisfaction with the skills of the trainers of the data outsourcing process	35
4.1.17	Respondents classification based on the most important module of data outsourcing	36
4.1.18	Respondents classification based on the most important feature of the data outsourcing	37
4.1.19	Respondents classification based on the algorithm used is best in the organization	38
4.1.20	Respondents classification based on the security algorithm can be suggested to the organization for the better usage	39

LIST OF CHARTS

CHART NO	TITLE	PAGE NO
4.1.1	Respondents classification based on Age	20
4.1.2	Respondents classification based on Gender	21
4.1.3	Respondents classification based on Educational Qualification	22
4.1.4	Respondents classification based on Experience in Information Security	23
4.1.5	Respondents classification based on Monthly Income (in Rs)	24
4.1.6	Respondents classification based on the algorithm that will be the best suggested	25
4.1.7	Respondents classification based on when the server will be usually upgraded in the company	26
4.1.8	Respondents classification based on starting to upload the confidential data	27
4.1.9	Respondents classification based on ensuring the security of the data in the organization	28
4.1.10	Respondents classification based on expecting the management to provide security of data	29
4.1.11	Respondents classification based on the algorithm that is presently used for secure data outsourcing	30
4.1.12	Respondents classification based on the satisfaction with the performance of the data outsourcing	31
4.1.13	Respondents classification based on the satisfaction with the depth of the working of the security algorithm	32
4.1.14	Respondents classification based on the satisfaction with the training methods of the	33

	working of the algorithm	
4.1.15	Respondents classification based on the satisfaction with the quality of training of the data outsourcing process	34
4.1.16	Respondents classification based on the satisfaction with the skills of the trainers of the data outsourcing process	35
4.1.17	Respondents classification based on the most important module of data outsourcing	36
4.1.18	Respondents classification based on the most important feature of the data outsourcing	37
4.1.19	Respondents classification based on the algorithm used is best in the organization	38
4.1.20	Respondents classification based on the security algorithm can be suggested to the organization for the better usage	39

CHAPTER – 1

INTRODUCTION

Secure Data Outsourcing

Recently, as the development of cloud computing, users' concerns about data security are the main obstacles that impede cloud computing from widely adopted. These concerns are originated from the fact that sensitive data resides in public cloud, which is maintained and operated by untrusted Cloud Service Provider (CSP). ABE (Attribute Based Encryption) provides a secure way that allows data owner to share outsourced data on untrusted storage server instead of trusted server with specified group of users. This advantage makes the methodology appealing in cloud storage that requires secure access control for a large number of users belonging to different organizations. Nevertheless, one of the main efficiency drawbacks of ABE is that the computational cost during decryption phase grows with the complexity of the access formula. Thus, before widely deployed, there is an increasing need to improve the efficiency of ABE. To address this problem, outsourced ABE, which provides a way to outsource intensive computing task during decryption to CSP without revealing data or private keys, was introduced. It has a wide range of applications. For example, in the mobile cloud computing consisting of mobile devices or sensors as information collection nodes, user terminal (e.g. mobile device) has limited computation ability to independently finish basic encryption or decryption to protect sensitive data residing in public cloud. Outsourced ABE allows user to perform heavy decryption through “borrowing” the computation resources from CSP. Therefore, with this paradigm, the computation/storage intensive tasks can be performed even by resource-constrained users. Beyond the heavy decryption outsourced, we observe that the attribute authority has to deal with a lot of heavy computation in a scalable system. More precisely, the attribute authority has to issue private keys to all users, but yet generation of private key typically requires large modular exponentiation computation, which grows linearly with the complexity of the predicate formula. When a large number of users call for their private keys, it may overload attribute authority.

Moreover, key management mechanism key revocation in particular, is necessary in a secure and scalable ABE system.

As organizations and individuals adopt cloud technologies, many have become aware of the serious concerns regarding security and privacy of accessing personal and confidential information over the Internet. In particular, the recent and continuing data breaches highlight the need for more secure cloud storage systems. While it is generally agreed that encryption is necessary, cloud providers often perform the encryption and maintain the private keys instead of the data owners. That is, the cloud can read any data it desired, providing no privacy to its users. The storage of private keys and encrypted data by the cloud provider is also problematic in case of data breach. Hence, researchers have actively been exploring solutions for secure storage on private and public clouds where private keys remain in the hands of data owners.

1.1 INDUSTRY PROFILE

According to National Association of Software and Service Companies (Nasscom), the Indian IT industry's revenue touched US\$ 227 billion in FY22, a 15.5% YoY growth and was estimated to have touched US\$ 245 billion in FY23.

In 2022, the Indian domestic IT & Business Services market was valued at US\$ 13.87 billion and recorded a 7.4% year-over-year (YoY) growth as compared to 7.2% in 2021.

IT spending in India is expected to increase to US\$ 110.3 billion in 2023 from an estimated US\$ 81.89 billion in 2021.

Indian software product industry is expected to reach US\$ 100 billion by 2025. Indian companies are focusing on investing internationally to expand their global footprint and enhance their global delivery centres.

The data annotation market in India stood at US\$ 250 million in FY20, of which the US market contributed 60% to the overall value. The market is expected to reach US\$ 7 billion by 2030 due to accelerated domestic demand for AI.

India's IT industry is likely to hit the US\$ 350 billion mark by 2026 and contribute 10% towards the country's gross domestic product (GDP), Infomerics Ratings said in a report.

As an estimate, India's IT export revenue rose by 9% in constant currency terms to US\$ 194 billion in FY23.

Market size of IT industry in India

Export of IT services has been the major contributor, accounting for more than 53% of total IT exports (including hardware).

BPM and engineering and R&D (ER&D) and software products exports accounted for 22% and 25%, respectively of total IT exports during FY23.

Exports from the Indian IT industry stood at US\$ 178 billion in FY22. Export of IT services was the major contributor, accounting for more than 51% of total IT export (including hardware). BPM and engineering and R&D (ER&D) and software products export accounted for 20.78% each to total IT exports during FY21. The ER&D market was expected to grow to US\$ 42 billion by 2022.

The IT industry added 2.9 lakh new jobs taking the industry's workforce tally to 5.4 million people in FY23.

INVESTMENTS/ DEVELOPMENTS

Indian IT's core competencies and strengths have attracted significant investment from major countries and companies.

- Direct employment in the IT services and BPO/ITeS segment was estimated to reach 5.4 million in FY23 with an addition of 290,000 people.
- The revenue of India's public cloud services market totalled US\$ 6.2 billion in 2022, and it is expected to reach US\$ 17.8 billion by 2027 growing at a CAGR of 23.4%.
- In November 2022, ICICI Bank introduced two new products for its NRI clients - Loan against Deposits (LAD) and Dollar Bonds, at its branch in GIFT City.
- In November 2022, Amazon Web Services announced the launch of its second AWS infrastructure region in India - the AWS Asia Pacific (Hyderabad) Region. By 2030, it is anticipated that the region will support more than 48,000 full-time jobs annually thanks to investments totaling more than US\$ 4.4 billion in India.
- In November 2022, Google established a partnership with local gaming startup SuperGaming through its Google Cloud division. As part of the collaboration, game developers who use Google Cloud to create, host, and distribute their games will have access to SuperGaming's SuperPlatform game engine.
- HDFC Bank partnered with Flywire to enable their customers to make fee payments digitally to overseas colleges and universities.
- In August 2022, Network People Services Technologies (NPST) announced that it was working on a banking super app. The high-end platform can be used by banks, fintech companies, and other BFSI players, and it will offer a seamless user experience with all banking, financial, and transactional services combined into a robust, smart app.
- In August 2022, PwC India announced that it was planning to hire 10,000 employees in the cloud and digital technologies space over the next five years.

- In October 2022, PE/VC investments in the technology sector stood at US\$ 157 million across 12 deals.
- The computer software and hardware sector in India attracted cumulative foreign investment (FDI) inflows worth US\$ 95.88 billion between April 2000-June 2023. The sector ranked second in FDI inflows as per the data released by Department for Promotion of Industry and Internal Trade (DPIIT). Computer software and hardware make up 15% of the cumulative FDI inflows.
- In July 2022, the Union Bank of India (UBI) launched a Metaverse Virtual Lounge and Open Banking Sandbox environment in partnership with Tech Mahindra.
- In June 2022, ZStack International, a worldwide market leader in cloud computing, IaaS and PaaS solutions, announced that they were entering into India and the SAARC Region.
- In June 2022, Redington India, an IT provider, entered into a multi-year strategic agreement with Amazon Web Services (AWS) to drive cloud technology adoption in India.
- American-Irish consumer credit reporting company Experian is planning a major expansion of its global innovation centre (GIC) in Hyderabad to about 4,000 employees over the next three to five years. According to sources, GIC will concentrate on employing emerging technologies in the BFSI (banking, financial services, and insurance) sector, including cloud computing, big data analytics, artificial intelligence, and machine learning.
- In 2021, PE investments in the IT sector stood at US\$ 23.4 billion.
- India's IT startup ecosystem received record investments of nearly US\$ 36 billion in privately held companies in 2021, up from US\$ 11 billion in 2020.
- In March 2022, Licious, India's largest tech-first, fresh animal protein brand, raised US\$ 150 million in a Series F2 funding round.
- In March 2022, Byju's raised US\$ 800 million in funding as part of a pre-IPO round, which values the Bengaluru-based company at about US\$ 22 billion.

- In March 2022, debt marketplace CredAvenue raised US\$ 137 million in a funding round led by Insight Partners, B Capital Group and Dragoneer Investment Group, which propelled the startup's valuation to US\$ 1.3 billion.
- In February 2022, Hasura, a software company which helps developers with tools, raised US\$ 100 million in a fresh funding round led by Greenoaks Capital, which elevate the company into a unicorn.
- In January 2022, Google announced plans to invest US\$ 1 billion in India's Bharti Airtel Ltd. to push India's digital ecosystem.
- Amazon partnered with Airtel to sell Amazon Web Services (AWS) to its customers and intends to inject US\$ 1.6 billion in their two upcoming data centres in Hyderabad.
- In November 2021, Wipro partnered with TEOCO to build solutions for communication service providers (CSPs) to improve network automation, efficiency, flexibility and reliability.
- In August 2021, Tata Consultancy Services was adjudged a leader in the NelsonHall NEAT for CX Services in Banking, Financial Services and Insurance (BFSI).

GOVERNMENT INITIATIVES

Some of the major initiatives taken by the government to promote the IT and ITeS sector in India are as follows:

- Cabinet approved PLI Scheme – 2.0 for IT Hardware with a budgetary outlay Rs. 17,000 crore (US\$ 2.06 billion).
- In the Union Budget 2023-24, the allocation for IT and telecom sector stood at Rs. 97,579.05 crore (US\$ 11.8 billion).
- In September 2022, the new Telecommunications Bill 2022 was published for public consultation by the Ministry of Communications as a move toward creating a new telecom framework in India.

- In August 2022, the Indian Computer Emergency Response Team (CERT-In), in collaboration with the Cyber Security Agency of Singapore (CSA), successfully planned and carried out the "Synergy" Cyber Security Exercise for 13 countries to build network resilience against ransomware attacks.
 - In June 2022, STPI Director General Mr. Arvind Kumar stated that exports through STPI units have increased from Rs. 17 crore (US\$ 2.14 million) in 1992 to Rs. 5.69 lakh crore (US\$ 71.65 billion) in 2022.
 - In May 2022, it was announced that Indians can now avail their Digilocker services through Whatsapp to get easy access to their official documents.
 - In April 2022, the Indian Computer Emergency Response Team (CERT-In) issued Directions to strengthen the cybersecurity in the country.
 - In the Union Budget 2022-23, the allocation for IT and telecom sector stood at Rs. 88,567.57 crore (US\$ 11.58 billion).
 - The government introduced the STP Scheme, which is a 100% export-oriented scheme for the development and export of computer software, including export of professional services using communication links or physical media.
 - In November 2021, the government launched the Internet Exchange in Uttarakhand to enhance the quality of internet services in the state.
 - The Karnataka government signed three MoUs worth US\$ 13.4 million (Rs. 100.52 crore) to help the state's emerging technology sector.
 - In September 2021, the Indian government announced a plan to build a cyber-lab for the 'Online Capacity Building Programme on Crime Investigation, Cyber Law and Digital Forensics' to strengthen cyber security capabilities.
-
- In September 2021, the Ministry of Electronics and Information Technology (MeitY) organised a workshop under the theme of 'Connecting all Indians' to promote public and private stakeholders' interest in the country and expand internet access to remote areas.
 - In September 2021, the Indian government launched the Meghalaya Enterprise Architecture Project (MeghEA) to boost service delivery and governance in the

state by leveraging digital technologies, to make Meghalaya a high-income state by 2030.

- In September 2021, the Indian government launched Phase II of Visvesvaraya PhD Scheme to encourage research in 42 emerging technologies in information technology (IT), electronics system design & manufacturing (ESDM) and information technology enabled services (ITES).
- In September 2021, the Indian government inaugurated five National Institute of Electronics & Information Technology (NIELIT) Centres in three North-Eastern states to boost availability of training centres and employment opportunities.
- On July 2, 2021, the Ministry of Heavy Industries and Public Enterprises launched six technology innovation platforms to develop technologies for globally competitive manufacturing in India. The six technology platforms have been developed by IIT Madras, Central Manufacturing Technology Institute (CMTI), International Centre for Automotive Technology (iCAT), Automotive Research Association of India (ARAI), BHEL, and HMT, in association with IISc Bangalore.

The Department of Telecom, Government of India and Ministry of Communications, Government of Japan, signed an MoU to enhance cooperation in areas of 5G technologies, telecom security and submarine optical fibre cable system.

ROAD AHEAD

India is the topmost offshoring destination for IT companies across the world. Having proven its capabilities in delivering both on-shore and off-shore services to global clients, emerging technologies now offer an entire new gamut of opportunities for top IT firms in India.

The Indian IT & business services industry is expected to grow to US\$ 19.93 billion by 2025. Spending on information technology in India is expected to reach US\$ 144 billion

in 2023. By 2026, widespread cloud utilisation can provide employment opportunities to 14 million people and add US\$ 380 billion to India's GDP.

As per a survey by Amazon Web Services (2021), India is expected to have nine times more digitally skilled workers by 2025.

IT spending in India is expected to increase to US\$ 110.3 billion in 2023 from an estimated US\$ 81.89 billion in 2021.

In November 2021, Mr. Piyush Goyal, Minister of Commerce and Industry, Consumer Affairs, Food and Public Distribution and Textiles, lauded the Indian IT sector for excelling its competitive strength with zero government interference. He further added that service exports from India has the potential to reach US\$ 1 trillion by 2030.

1.2 Concepts

The concept of the project is represented using the modules involved and review of literature in this project.

Modules

- User Module
- Owner Module
- Authority Grant Module
- User Revocation Module

➤ File Access Module

User Module

In this module, the resources that are performing in this project are created. The user will be created by making the user to fill a registration form. On successful submission of the registration form, the user can login using the credentials. The data owner will be the user who can upload the data into the cloud that can be used by the users.

Owner Module

The new file creation modules, enables the owner to upload the contents from the cloud server. The content that is about to be uploaded will have various attributes like file name, file description and the key information. After providing the data of the attributes, the uploading of the content will be successful.

Authority Grant Module

In this module, the authority for the users and the owners will be created. The user will be provided with the access rights like read, write privileges by the owner. The cryptographic techniques like key generation, encryption will be implemented in this module. Based on the privilege set to the users, the users can access the data. The admin manages the user and the data details.

User Revocation Module

In this module, a system just maintains some state information of users' keys and assigns new value for expiration time to a user's key when updating it. When re-encrypting data

files, the data owner just needs two exponentiations for cipher text components associated with the attribute.

File Access Module

In this operation, we discuss the decrypting operation of encrypted data files. A user first obtains with the decrypt algorithm and then decrypts data files using DEKs. The cost of decrypting a ciphertext varies depending on the key used for decryption. Even for a given key, the way to satisfy the associated access tree may be various. The algorithm consists of two pairing operations for every leaf node used to satisfy the tree, one pairing for each translating node on the path from the leaf node used to the root and one exponentiation for each node on the path from the leaf node to the root. The techniques like key matching, decryption will be performed in this module. So the computation complexity varies depending on the access tree and key structure. It should be noted that the decryption is performed at the data consumers.

AES ALGORITHM

Steps

1. Derive the set of round keys from the cipher key.
2. Initialize the state array with the block data (plaintext).
3. Add the initial round key to the starting state array.
4. Perform nine rounds of state manipulation.
5. Perform the tenth and final round of state manipulation.
6. Copy the final state array out as the encrypted data (ciphertext).

1.3 Need for the study

The need for the study is to know that in the current business world, the use of online data sharing is growing day by day. The data may be a highly confidential data. This data should be delivered only to the concerned destination. During the data transfer, the data should be protected from the hackers.

1.4 Problem of the study

The problem of the project is based on the need to understand that each user will have a functionality to be performed. The data owner may have to upload the data and the consumer may need to download the data. In the field of information security management, it is very essential to provide security and privacy to both the users. This project may make use of cryptographic techniques by means of algorithms available for providing security and privacy. The system should be very secure and provide privacy to the users.

1.5 Objectives of the study

The objectives of the study are based on the primary and secondary objectives.

Primary Objective

The primary objective of the study is to understand the secure outsourcing practices performed in the company.

Secondary Objectives

The secondary objective of the study is

- To understand the concept of data outsourcing.
- To identify the users involved, ensure security and privacy to the users.
- To implement the cryptography algorithm to achieve the security of the data.

CHAPTER - 2

LITERATURE REVIEW

2.1 Review of Literature

According to **D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano**, “Public key encryption with keyword search”, 2004, this paper studies the problem of searching on data that is encrypted using a public key system. Consider user Bob who sends email to user Alice encrypted under Alice's public key. An email gateway wants to test whether the email contains the keyword “urgent” so that it could route the email accordingly. Alice, on the other hand does not wish to give the gateway the ability to decrypt all her messages. We define and construct a mechanism that enables Alice to provide a key to the gateway that enables the gateway to test whether the word “urgent” is a keyword in the email without learning anything else about the email. We refer to this mechanism as Public Key Encryption with keyword Search. As another example, consider a mail server that stores various messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some specific keyword, but learn nothing else. We define the concept of public key encryption with keyword search and give several constructions.

According to **B. Waters, D. Balfanz, G. Durfee, and D. K. Smetters**, “Building an encrypted and searchable audit log”, 2004, audit log is important part of any software system. These audit logs may contain sensitive information, poses a threat to privacy and information security, so should be prohibited against any illegal reading and alteration or deletion. The best way to do this is encryption. The key challenges in an encrypted audit log are speed of log and search process, correctness of query and relevance of search results and log size. We describe an approach for constructing searchable encrypted audit logs which can be combined with any number of existing approaches for creating tamper-

resistant logs. Even though previous searchable encryption schemes allow users to search encrypted data by keywords securely, these techniques only support exact keyword search and will fail if there are some spelling errors or if some morphological variants of words are used. In this paper, we provide the solution for fuzzy keyword search over encrypted data. K-grams are used to produce fuzzy results. Our technique for keyword search on encrypted data has wide application beyond searchable audit logs.

According to **M. Ding, F. Gao, Z. Jin, and H. Zhang**, “An efficient public key encryption with conjunctive keyword search scheme based on pairings”, 2012, the Public Key Encryption with Conjunctive Keyword Search (PECK) scheme enables one to search a document included multiple encrypted keywords without compromising any original data information. The existing PECK schemes mostly depend on pairings and authenticated channel to achieve searchable encryption. In this paper, we propose a new PECK scheme based on pairings, where no pairing operations are involved in the encryption and trapdoor phases and no secure channel is needed between server and users. In comparison with previous schemes, our scheme can achieve high efficiencies in both computation and communication. And then we further discuss on the consistency, efficiency and security of the proposed scheme.

According to **F. Kerschbaum**, “Secure conjunctive keyword searches for unstructured text”, 2011, this paper studies the setting in which a user stores encrypted documents (e.g. e-mails) on an untrusted server. In order to retrieve documents satisfying a certain search criterion, the user gives the server a capability that allows the server to identify exactly those documents. Work in this area has largely focused on search criteria consisting of a single keyword. If the user is actually interested in documents containing each of several keywords (conjunctive keyword search) the user must either give the server capabilities for each of the keywords individually and rely on an intersection calculation (by either the server or the user) to determine the correct set of documents, or alternatively, the user may store additional information on the server to facilitate such searches. Neither solution is desirable; the former enables the server to learn which documents match each

individual keyword of the conjunctive search and the latter results in exponential storage if the user allows for searches on every set of keywords. We define a security model for conjunctive keyword search over encrypted data and present the first schemes for conducting such searches securely. We propose first a scheme for which the communication cost is linear in the number of documents, but that cost can be incurred offline before the conjunctive query is asked. We propose a second scheme whose communication cost is on the order of the number of keyword fields and whose security relies on a new hardness assumption.

According to **C. Hu and P. Liu**, “Public key encryption with ranked multikeyword search”, 2013, it is necessary for the cloud server to allow multi-keyword query and provide result ranking to effectively retrieve data from the large number of encrypted data stored in the cloud storage. A related work that supports privacy preserving ranked multi-keyword search over encrypted cloud data (MRSE) is not in public-key model. In this paper, the definition of an extension of PEKS called Public-Key Encryption with Ranked Multi-Keyword Search (PERMKS) is proposed, which means that the receiver could query any subset of keywords and the number of queried keywords appearing in the data can be provided which evaluate the similarity ranking of the data to the search query. Then, a construction of public-key encryption with ranked multikeyword search scheme with sub-linear cipher texts based on anonymous hierarchical identity-based encryption (AHIBE) is put forward and its security is analyzed.

According to **Z. Fu, X. Sun, N. Linge, and L. Zhou**, “Achieving effective cloud search services: multi-keyword ranked search over encrypted cloud data supporting synonym query”, 2014, in recent years, consumer-centric cloud computing paradigm has emerged as the development of smart electronic devices combined with the emerging cloud computing technologies. A variety of cloud services are delivered to the consumers with the premise that an effective and efficient cloud search service is achieved. For consumers, they want to find the most relevant products or data, which is highly desirable in the "pay-as-you use" cloud computing paradigm. As sensitive data (such as photo albums, emails, personal health records, financial records, etc.) are encrypted before

outsourcing to cloud, traditional keyword search techniques are useless. Meanwhile, existing search approaches over encrypted cloud data support only exact or fuzzy keyword search, but not semantics-based multi-keyword ranked search. Therefore, how to enable an effective searchable system with support of ranked search remains a very challenging problem. This paper proposes an effective approach to solve the problem of multi-keyword ranked search over encrypted cloud data supporting synonym queries. The main contribution of this paper is summarized in two aspects: multi-keyword ranked search to achieve more accurate search results and synonym-based search to support synonym queries. Extensive experiments on real-world dataset were performed to validate the approach, showing that the proposed solution is very effective and efficient for multikeyword ranked searching in a cloud environment.

According to **C. L. A. Clarke, G. V. Cormack, and E. A. Tudhope**, “Relevance ranking for one to three term queries”, 2000, this paper investigates the application of a novel relevance ranking technique, cover density ranking, to the requirements of Web-based information retrieval, where a typical query consists of a few search terms and a typical result consists of a page indicating several potentially relevant documents. Traditional ranking methods for information retrieval, based on term and inverse document frequencies, have often been found to work poorly in this context. Under the cover density measure, ranking is based on term proximity and co-occurrence. Experimental comparisons show retrieval performance that compares favourably with previous work.

According to **H. Tuo and M. Wenping**, “An effective fuzzy keyword search scheme in cloud computing”, 2013, with the development of cloud computing, individuals and enterprises stored their data in the cloud for great flexibility and economic savings. Many keyword search schemes over encrypted data with privacy preserving were proposed, however, few schemes could efficiently support the similarity search. In this paper, we investigate the issue on fuzzy search over cloud data, then by using technique of filter, we

improve a efficient keyword search scheme to achieve fuzzy searching with low cost, which is suit for practical cloud computing.

According to **M. Zheng and H. Zhou**, “An efficient attack on a fuzzy keyword search scheme over encrypted data”, 2013, cipher text search ability is a promising method to securely store and retrieve outsourced data, like in secure cloud storage. But it is still hard to do fuzzy keyword search over outsourced cipher texts. In INFOCOM 2010, Li et al. proposed a fuzzy keyword search scheme over encrypted data based on edit distance. However, we find that their scheme is insecure. In this paper, we construct a special adversary to break the provable security reduction of their scheme. In theory, the main reason of the insecurity is caused by the independency of the indexes generated for different keywords. So an adversary can break the provable security according to some possessed messages learned from history interaction in the provable security.

According to **S. Zittrower and C. C. Zou**, “Encrypted phrase searching in the cloud”, 2012, as cloud computing is increasing in popularity, it is difficult to both maintain privacy in datasets while still providing adequate retrieval and searching procedures. This paper introduces a novel approach in the field of encrypted searching that allows both encrypted phrase searches and proximity ranked multi-keyword searches to encrypted datasets on untrusted cloud. By storing encrypted keyword-location data along with specially truncated encrypted keyword indexes in a relational database, we are able to allow for a full range of search features in our encrypted searches, something that has never been accomplished before. Furthermore, our approach permits the encrypted corpus and index to both be stored on cloud data servers. We modify currently available open-source search engine software to complete a prototype and provide results from experiments on a large scale real-world dataset that has more than half a million documents.

CHAPTER – 3

RESEARCH METHODOLOGY

Research Methodology

3.1 Research Approach

The questionnaire is prepared on the basis of objectives of the study. Direct contract is used for survey, contacting employees directly in order to collect data.

Sampling Area

The study is conducted for the employees of the industry.

Sampling Design

The researcher has used probability sampling in which simple random sampling is used.

Sample Size

50 Respondents are used for this study.

3.2 Collection of Data

The collection of the data is through the primary and secondary data.

Ø Primary Data

The primary data are those that are collected through questionnaire and direct

personal interview. The questionnaire was farmed in such a manner to obtain the correct information, graded suitably for the study. All the questionnaires were collected through personal contact from the respondents.

Ø Secondary Data

It is collected form company records, websites, documents, newspapers, journals, magazines, etc.

Expected Outcome

- Quality of Data
- Security of Data
- Privacy of Data

3.3 Limitations

- This study is limited to Chennai region only
- Since this is a major topic, time is a major limitation

CHAPTER – 4

DATA ANALYSIS

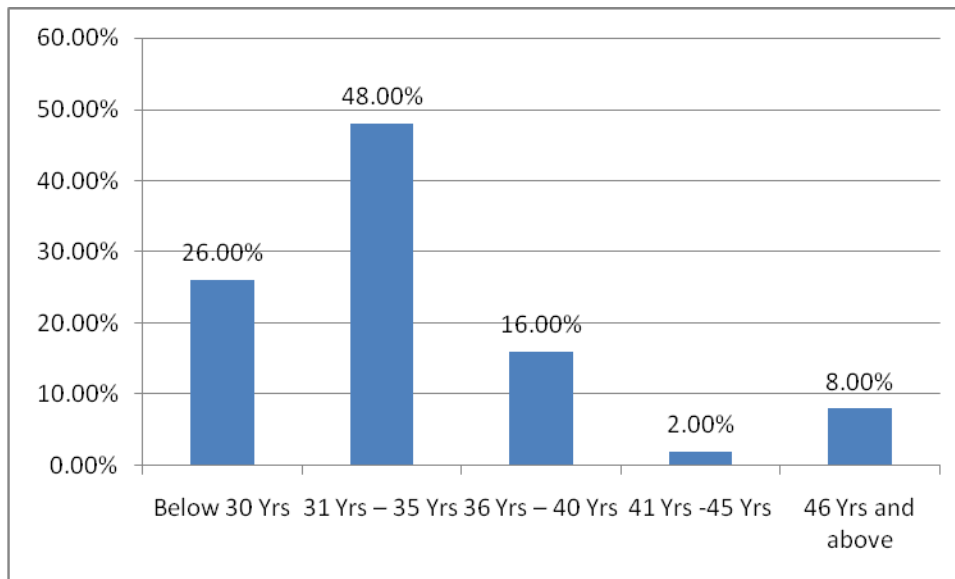
4.1 TABLES AND CHARTS

TABLE NO 4.1.1

Respondents classification based on Age

SI.No	Description	Frequency	Percentage
1	Below 30 Yrs	13	26.00%
2	31 Yrs – 35 Yrs	24	48.00%
3	36 Yrs – 40 Yrs	8	16.00%
4	41 Yrs -45 Yrs	1	2.00%
5	46 Yrs and above	4	8.00%
	Total	50	100%

CHART NO 4.1.1



Interpretation

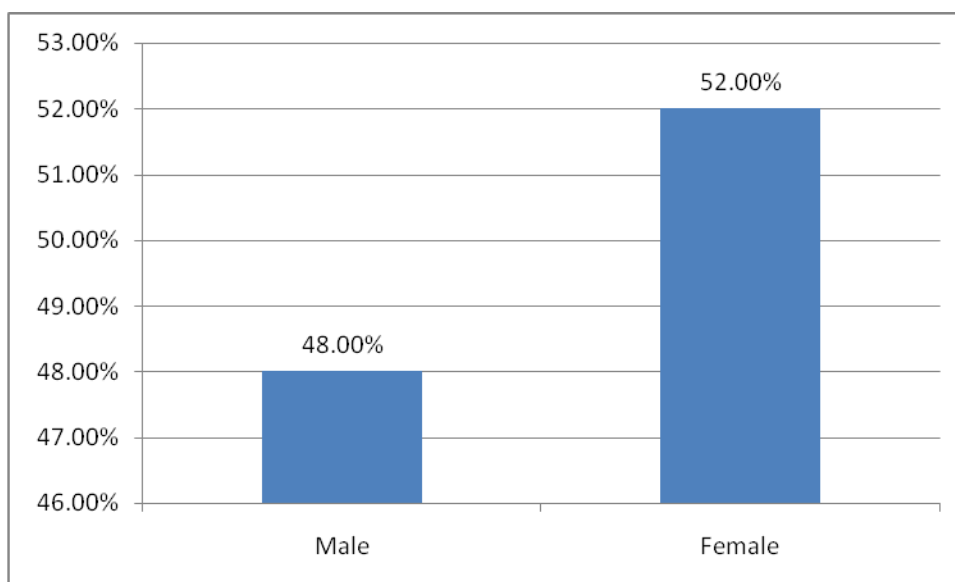
From the above table and chart, it is interpreted that 48.00% of the respondents are between 31-35 yrs, 26.00% of the respondents are below 30 Yrs, 16.00% of the respondents are between 36-40 yrs and 8.00% of the respondents are above 46 yrs and above.

TABLE NO 4.1.2

Respondents classification based on Gender

SI.No	Description	Frequency	Percentage
1	Male	24	48.00%
2	Female	26	52.00%
	Total	50	100%

CHART NO 4.1.2



Interpretation

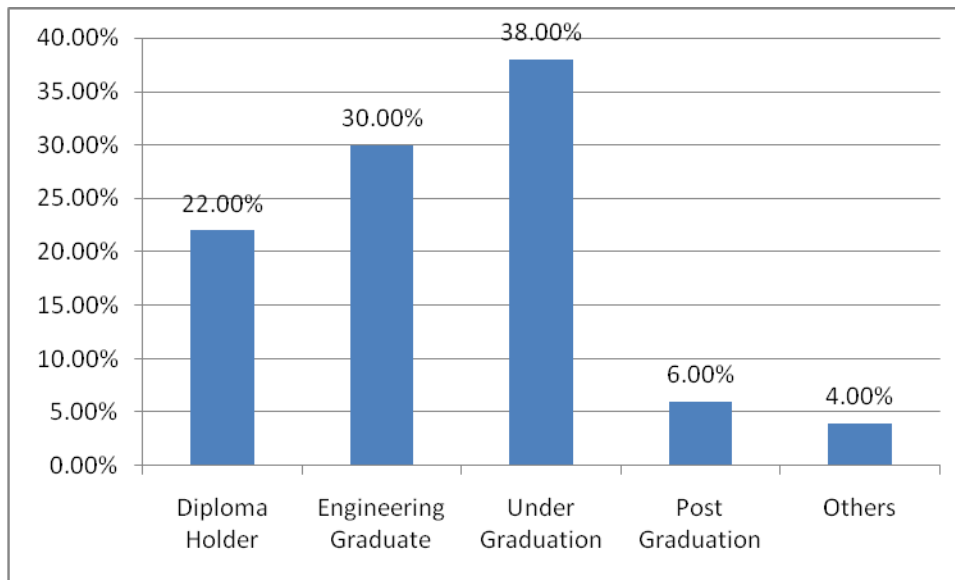
From the above table and chart, it is interpreted that 52.00% of the respondents are females and 48.00% of the respondents are male respectively.

TABLE NO 4.1.3

Respondents classification based on Educational Qualification

Sl.No	Description	Frequency	Percentage
1	Diploma Holder	11	22.00%
2	Engineering Graduate	15	30.00%
3	Under Graduation	19	38.00%
4	Post Graduation	3	6.00%
5	Others	2	4.00%
	Total	50	100%

CHART NO 4.1.3



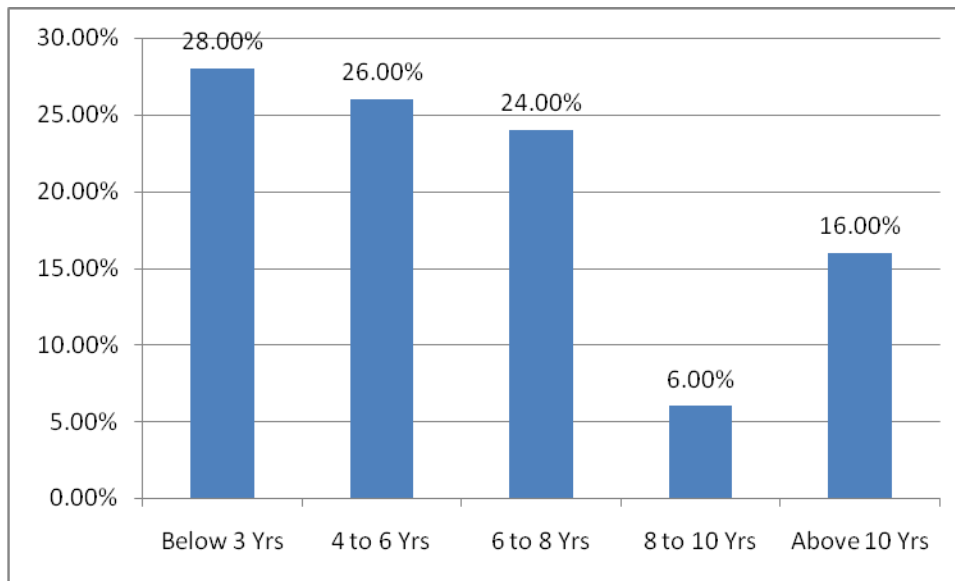
Interpretation

From the above table and chart, it is interpreted that 38.00% of the respondents are undergraduates, 30.00% of the respondents are engineering graduates, 22.00% of the respondents are diploma holders and 6.00% of the respondents are post graduates.

TABLE NO 4.1.4

Respondents classification based on Experience in Information Security

Sl.No	Description	Frequency	Percentage
1	Below 3 Yrs	14	28.00%
2	4 to 6 Yrs	13	26.00%
3	6 to 8 Yrs	12	24.00%
4	8 to 10 Yrs	3	6.00%
5	Above 10 Yrs	8	16.00%
	Total	50	100%

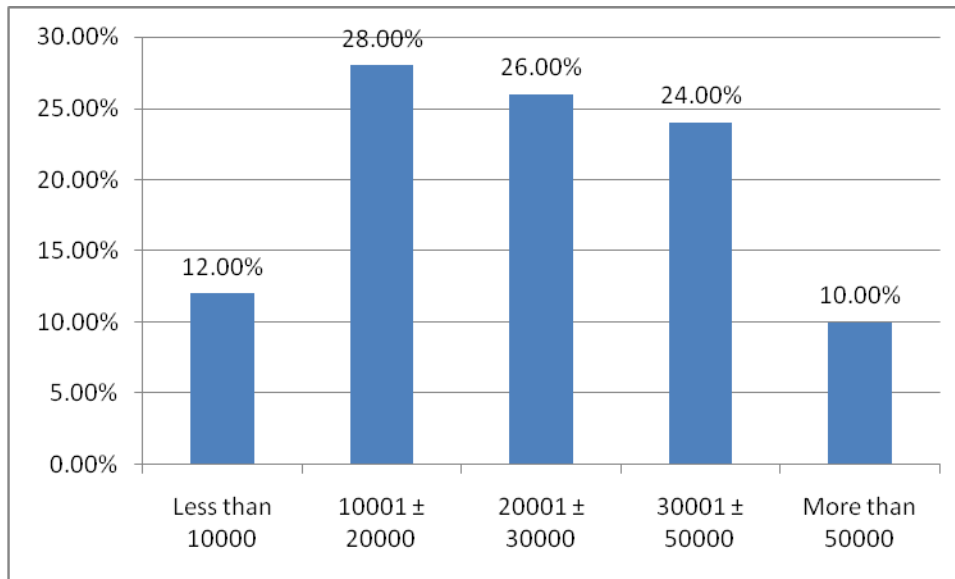
CHART NO 4.1.4**Interpretation**

From the above table and chart, it is interpreted that 28.00% of the respondents are below 3 years, 26.00% of the respondents are between 4 to 6 yrs, 24.00% of the respondents are between 6 to 8 yrs and 16.00% of the respondents are above 10 yrs of experience in information security.

TABLE NO 4.1.5

Respondents classification based on Monthly Income (in Rs)

Sl.No	Description	Frequency	Percentage
1	Less than 10000	6	12.00%
2	10001 ± 20000	14	28.00%
3	20001 ± 30000	13	26.00%
4	30001 ± 50000	12	24.00%
5	More than 50000	5	10.00%
	Total	50	100%

CHART NO 4.1.5**Interpretation**

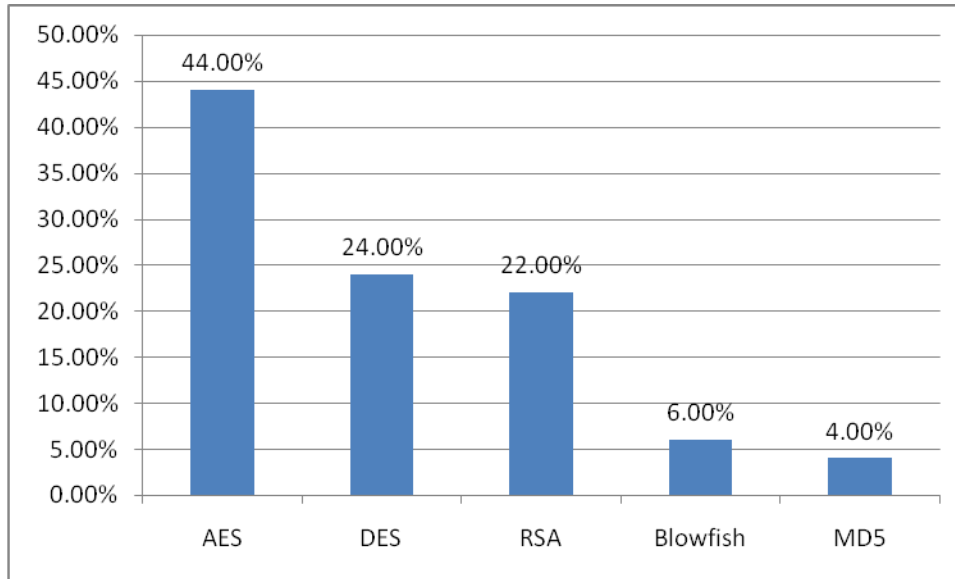
From the above table and chart, it is interpreted that 28.00% of the respondents are getting monthly income as 10001-20000, 26.00% of the respondents are getting monthly income as 20001-30000, 24.00% of the respondents are getting monthly income as 30001-50000, 10.00% of the respondents as 50001-60000 and 12.00% of the respondents are getting monthly income less than above 10000.

TABLE NO 4.1.6

Respondents classification based on the algorithm that will be the best suggested

SI.No	Description	Frequency	Percentage
1	AES	22	44.00%
2	DES	12	24.00%
3	RSA	11	22.00%
4	Blowfish	3	6.00%
5	MD5	2	4.00%
	Total	50	100%

CHART NO 4.1.6



Interpretation

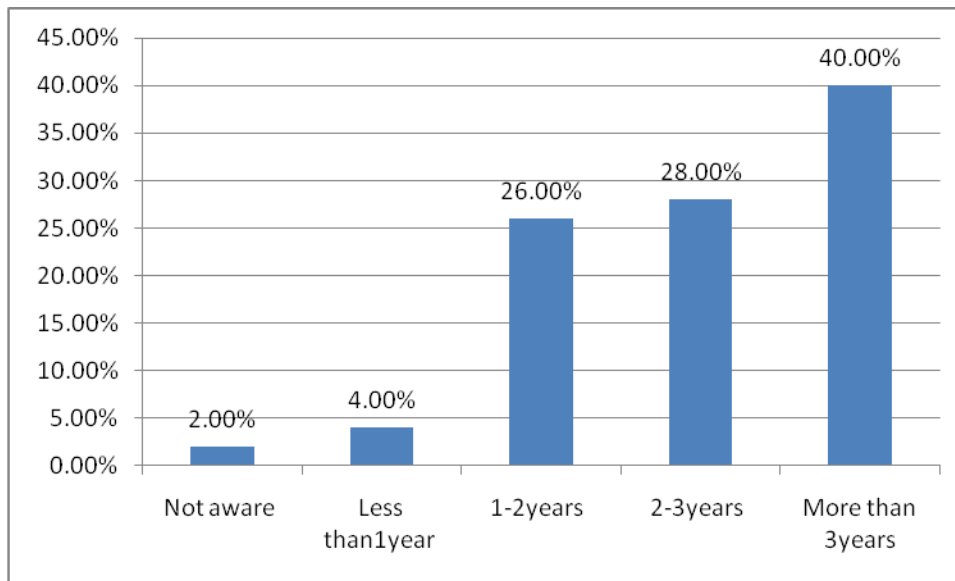
From the above table and chart, it is interpreted that 44.00% of the respondents as AES, 24.00% of the respondents as DES, 22.00% of the respondents as RSA, 6.00% of the respondents as Blowfish and 4.00% of the respondents as MD5 as the algorithm that will be the best suggested.

TABLE NO 4.1.7

Respondents classification based on when the server will be usually upgraded in the company

SI.No	Description	Frequency	Percentage
1	Not aware	1	2.00%
2	Less than1year	2	4.00%
3	1-2years	13	26.00%
4	2-3years	14	28.00%
5	More than 3years	20	40.00%
	Total	50	100%

CHART NO 4.1.7



Interpretation

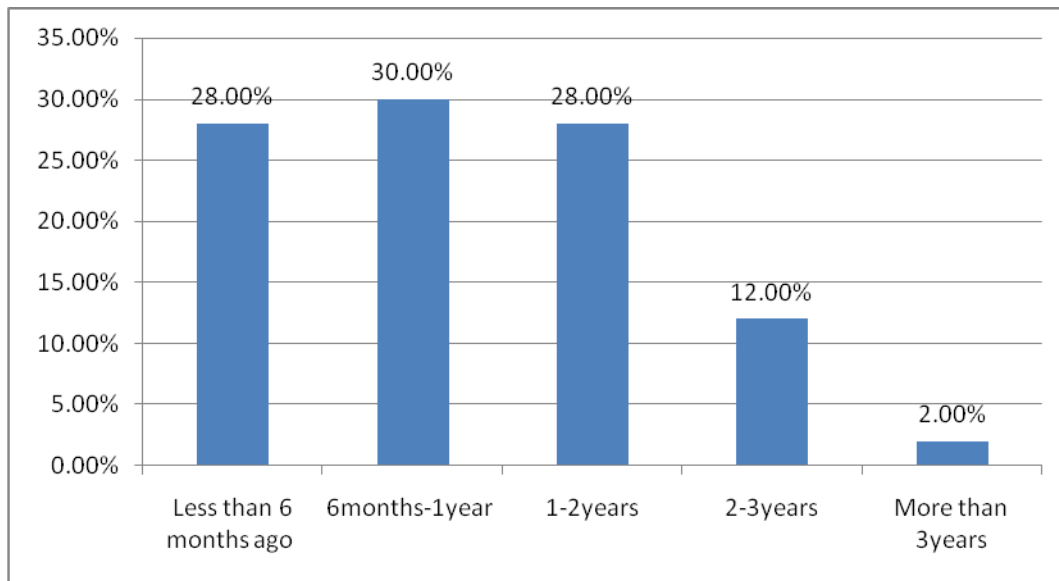
From the above table and chart, it is interpreted that 40.00% of the respondents as more than 3 years, 28.00% of the respondents as 2-3years, 26.00% of the respondents as 1-2years, 4.00% of the respondents as less than 1 year and 2.00% of the respondents are not aware based on when the server will be usually upgraded in the company.

TABLE NO 4.1.8

Respondents classification based on starting to upload the confidential data

SI.No	Description	Frequency	Percentage
1	Less than 6 months ago	14	28.00%
2	6months-1year	15	30.00%
3	1-2years	14	28.00%
4	2-3years	6	12.00%
5	More than 3years	1	2.00%
	Total	50	100%

CHART NO 4.1.8



Interpretation

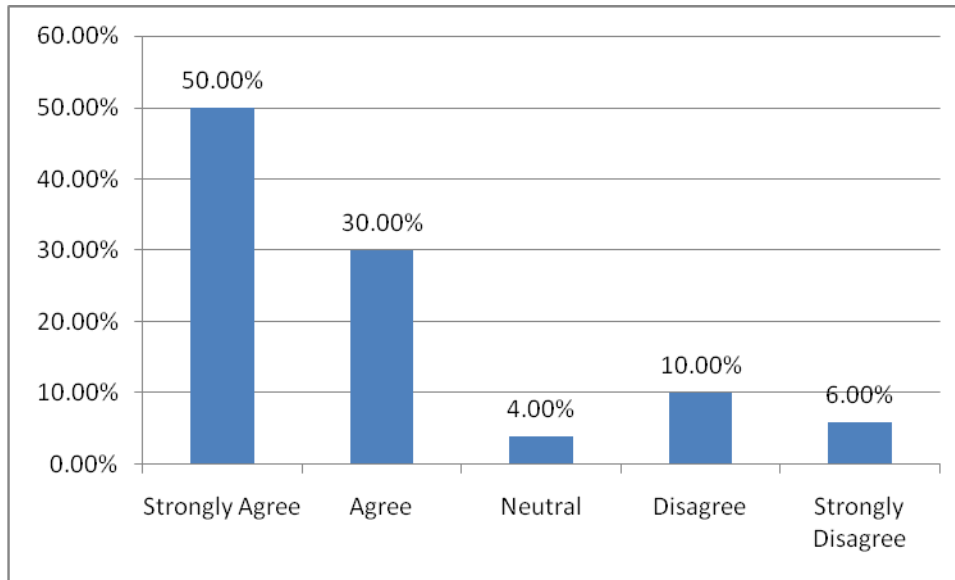
From the above table and chart, it is interpreted that 30.00% of the respondents as 6 months – 1 year, 28.00% of the respondents as less than 6months and 1-2years, 12.00% of the respondents as 2-3 years and 2.00% of the respondents as more than 3 years based on starting to upload the confidential data.

TABLE NO 4.1.9

Respondents classification based on ensuring the security of the data in the organization

Sl.No	Description	Frequency	Percentage
1	Strongly Agree	25	50.00%
2	Agree	15	30.00%
3	Neutral	2	4.00%
4	Disagree	5	10.00%
5	Strongly Disagree	3	6.00%
	Total	50	100%

CHART NO 4.1.9



Interpretation

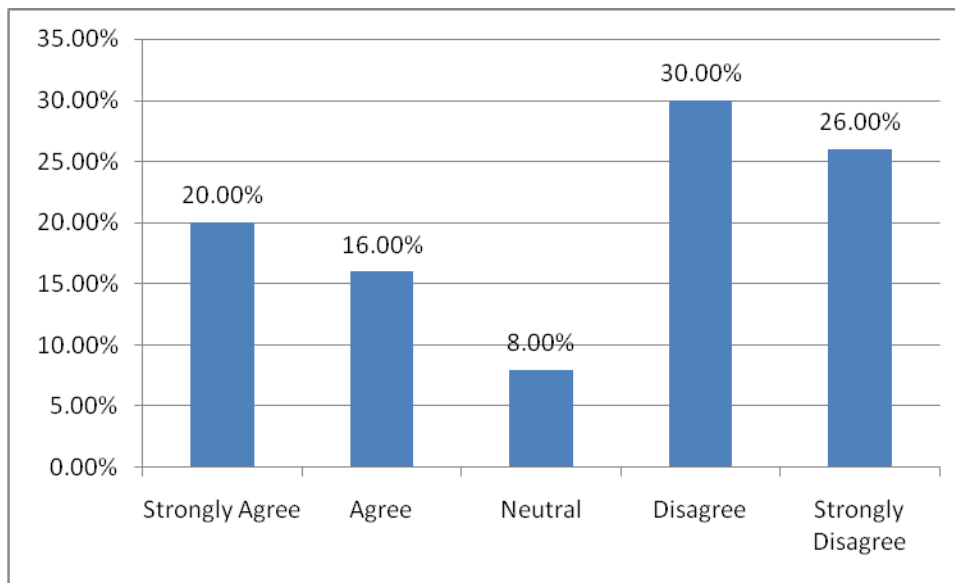
From the above table and chart, it is interpreted that 50.00% of the respondents strongly agree, 30.00% of the respondent agree, 10.00% of the respondent disagree, 6.00% of the respondent strongly disagree and 4.00% of the respondent are neutral based on ensuring the security of the data in the organization.

TABLE NO 4.1.10

Respondents classification based on expecting the management to provide security of data

SI.No	Description	Frequency	Percentage
1	Strongly Agree	10	20.00%
2	Agree	8	16.00%
3	Neutral	4	8.00%
4	Disagree	15	30.00%
5	Strongly Disagree	13	26.00%
	Total	50	100%

CHART NO 4.1.10



Interpretation

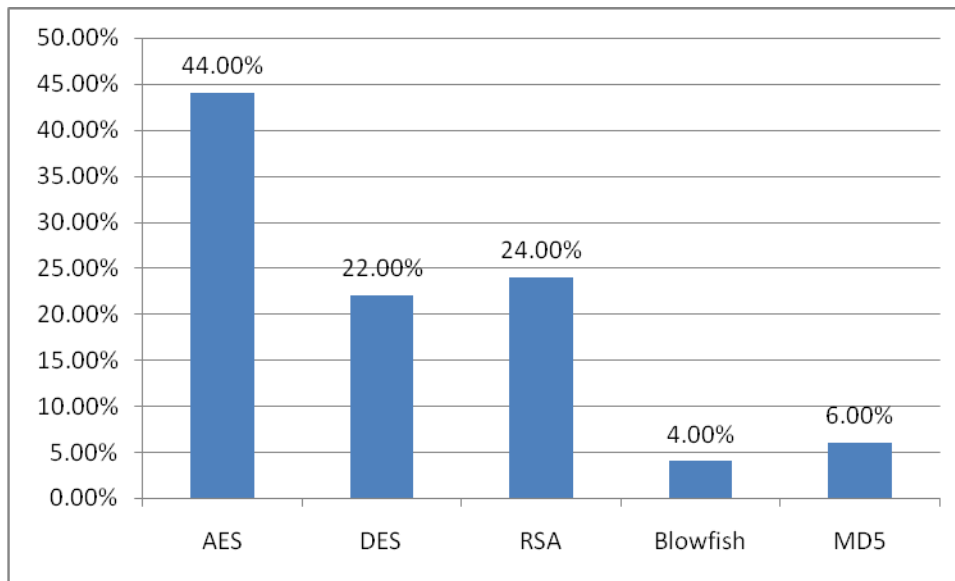
From the above table and chart, it is interpreted that 30.00% of the respondents disagree, 26.00% of the respondents strongly disagree, 20.00% of the respondents strongly agree, 16.00% of the respondents agree and 8.00% of the respondents are neutral based on expecting the management to provide security of data.

TABLE NO 4.1.11

Respondents classification based on the algorithm that is presently used for secure data outsourcing

Sl.No	Description	Frequency	Percentage
1	AES	22	44.00%
2	DES	11	22.00%
3	RSA	12	24.00%
4	Blowfish	2	4.00%
5	MD5	3	6.00%
	Total	50	100%

CHART NO 4.1.11



Interpretation

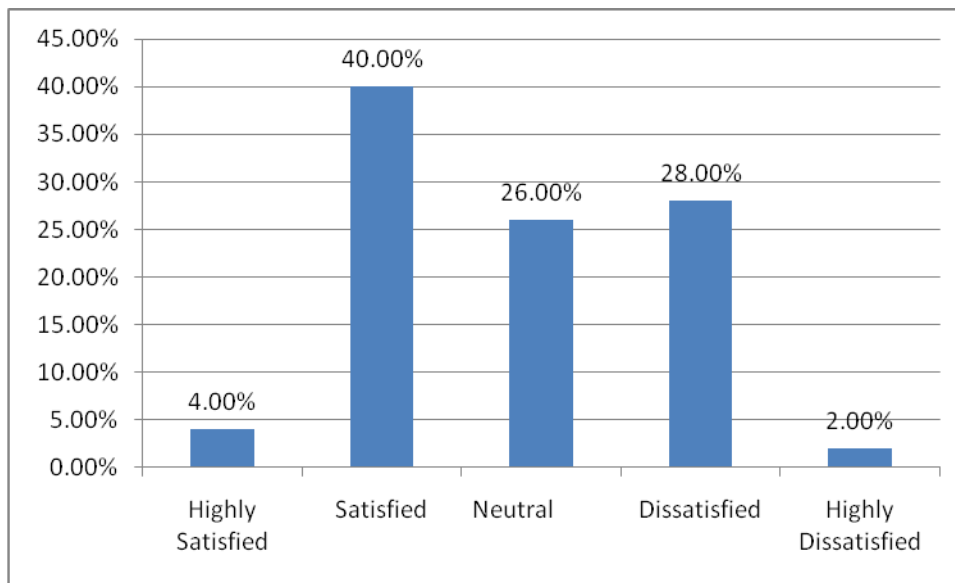
From the above table and chart, it is interpreted that 44.00% of the respondents as AES, 24.00% of the respondents as RSA, 22.00% of the respondents as DES, 6.00% of the respondents as MD5 and 4.00% of the respondents as blowfish based on the algorithm that is presently used for secure data outsourcing.

TABLE NO 4.1.12

Respondents classification based on the satisfaction with the performance of the data outsourcing

SI.No	Description	Frequency	Percentage
1	Highly Satisfied	2	4.00%
2	Satisfied	20	40.00%
3	Neutral	13	26.00%
4	Dissatisfied	14	28.00%
5	Highly Dissatisfied	1	2.00%
	Total	50	100%

CHART NO 4.1.12



Interpretation

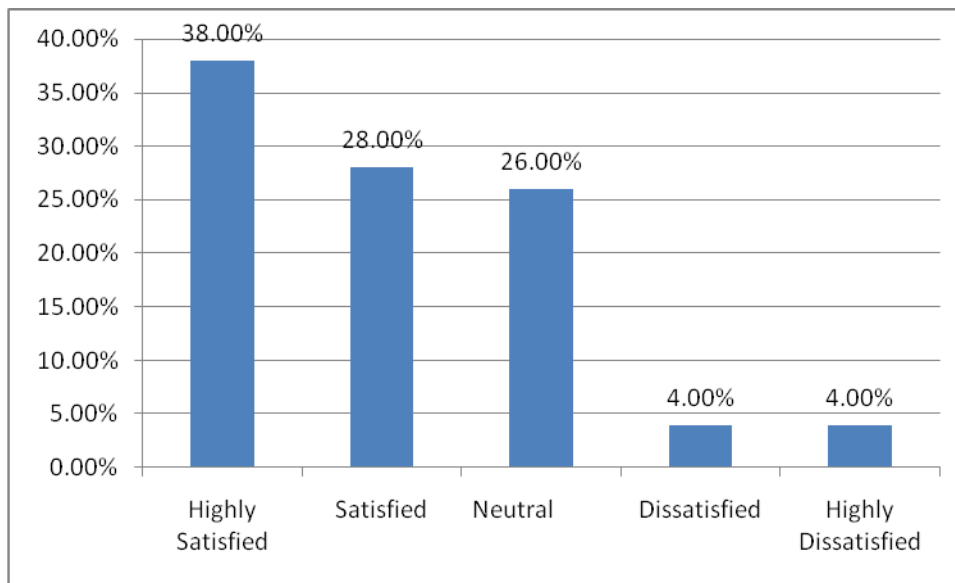
From the above table and chart, it is interpreted that 40.00% of the respondents are satisfied, 28.00% of the respondents are dissatisfied, 26.00% of the respondents are neutral, 4.00% of the respondents are highly satisfied and 2.00% of the respondents are highly dissatisfied based on the satisfaction with the performance of the data outsourcing.

TABLE NO 4.1.13

Respondents classification based on the satisfaction with the depth of the working of the security algorithm

Sl.No	Description	Frequency	Percentage
1	Highly Satisfied	19	38.00%
2	Satisfied	14	28.00%
3	Neutral	13	26.00%
4	Dissatisfied	2	4.00%
5	Highly Dissatisfied	2	4.00%
	Total	50	100%

CHART NO 4.1.13



Interpretation

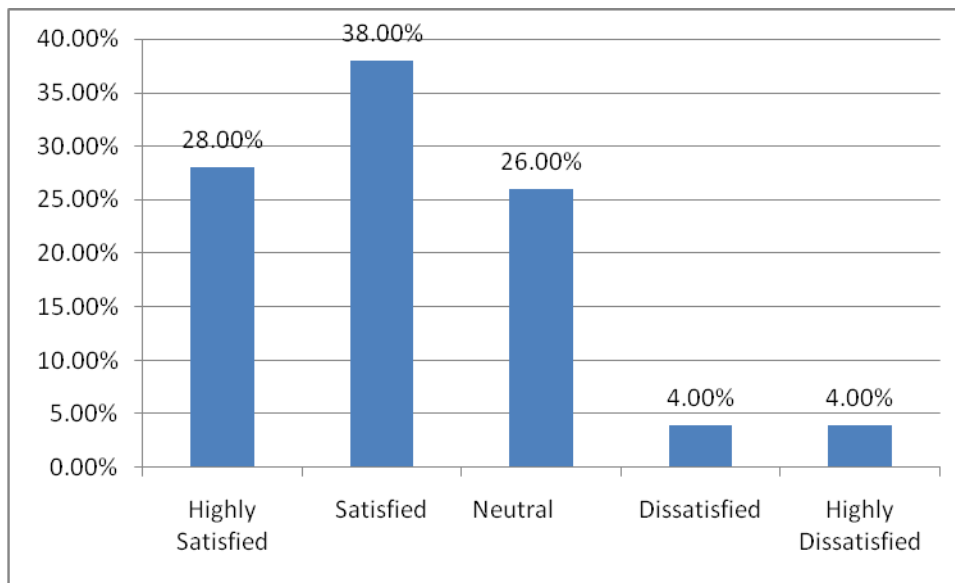
From the above table and chart, it is interpreted that 38.00% of the respondents are highly satisfied, 28.00% of the respondents are satisfied, 26.00% of the respondents are neutral and 4.00% of the respondents are dissatisfied and highly dissatisfied based on the satisfaction with the depth of the working of the security algorithm.

TABLE NO 4.1.14

Respondents classification based on the satisfaction with the training methods of the working of the algorithm

Sl.No	Description	Frequency	Percentage
1	Highly Satisfied	14	28.00%
2	Satisfied	19	38.00%
3	Neutral	13	26.00%
4	Dissatisfied	2	4.00%
5	Highly Dissatisfied	2	4.00%
	Total	50	100%

CHART NO 4.1.14



Interpretation

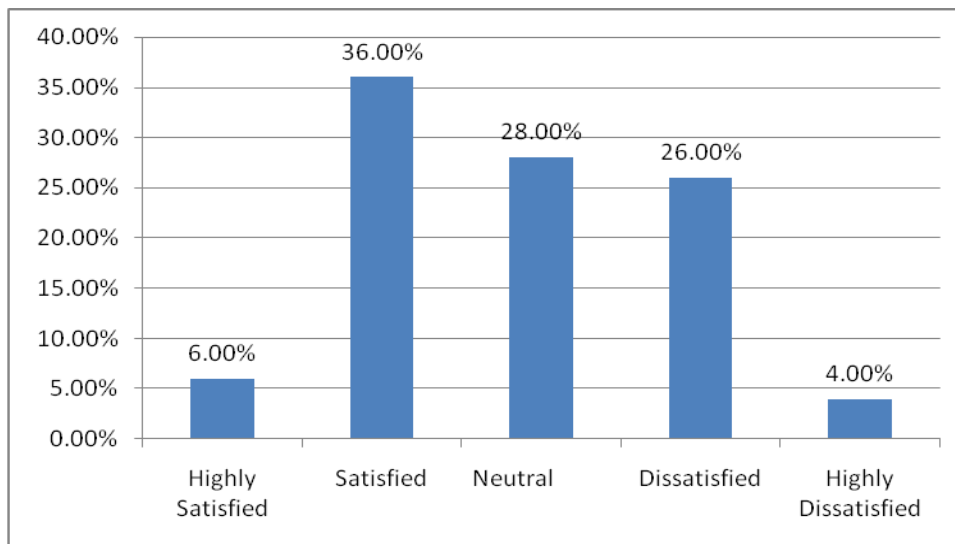
From the above table and chart, it is interpreted that 38.00% of the respondents are satisfied, 28.00% of the respondents are highly satisfied, 26.00% of the respondents are neutral and 4.00% of the respondents are dissatisfied and highly dissatisfied based on the satisfaction with the training methods of the working of the algorithm.

TABLE NO 4.1.15

Respondents classification based on the satisfaction with the quality of training of the data outsourcing process

SI.No	Description	Frequency	Percentage
1	Highly Satisfied	3	6.00%
2	Satisfied	18	36.00%
3	Neutral	14	28.00%
4	Dissatisfied	13	26.00%
5	Highly Dissatisfied	2	4.00%
	Total	50	100%

CHART NO 4.1.15



Interpretation

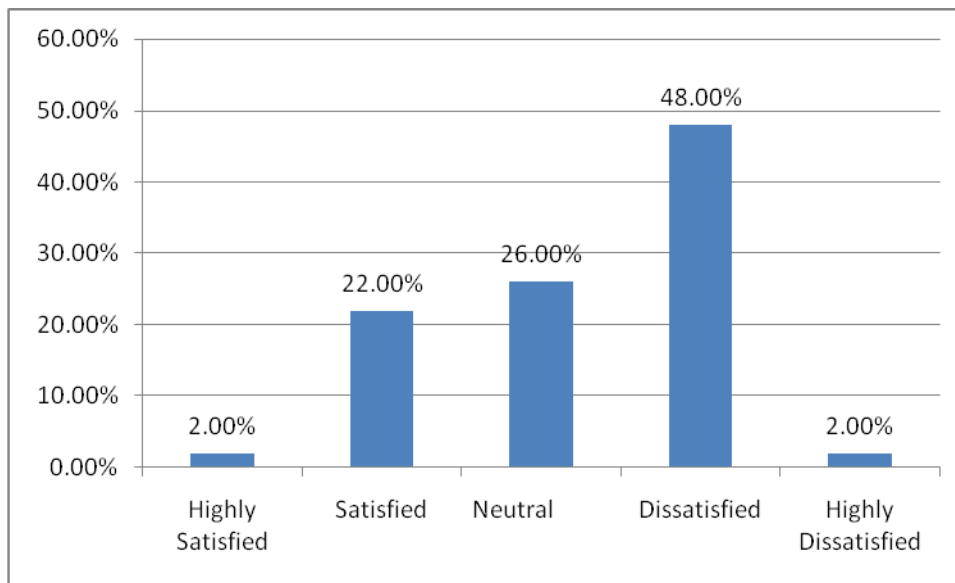
From the above table and chart, it is interpreted that 36.00% of the respondents are satisfied, 28.00% of the respondents are neutral, 26.00% of the respondents are dissatisfied, 6.00% of the respondents are highly satisfied and 4.00% of the respondents are highly dissatisfied based on the satisfaction with the quality of training of the data outsourcing process.

TABLE NO 4.1.16

Respondents classification based on the satisfaction with the skills of the trainers of the data outsourcing process

Sl.No	Description	Frequency	Percentage
1	Highly Satisfied	1	2.00%
2	Satisfied	11	22.00%
3	Neutral	13	26.00%
4	Dissatisfied	24	48.00%
5	Highly Dissatisfied	1	2.00%
	Total	50	100%

CHART NO 4.1.16



Interpretation

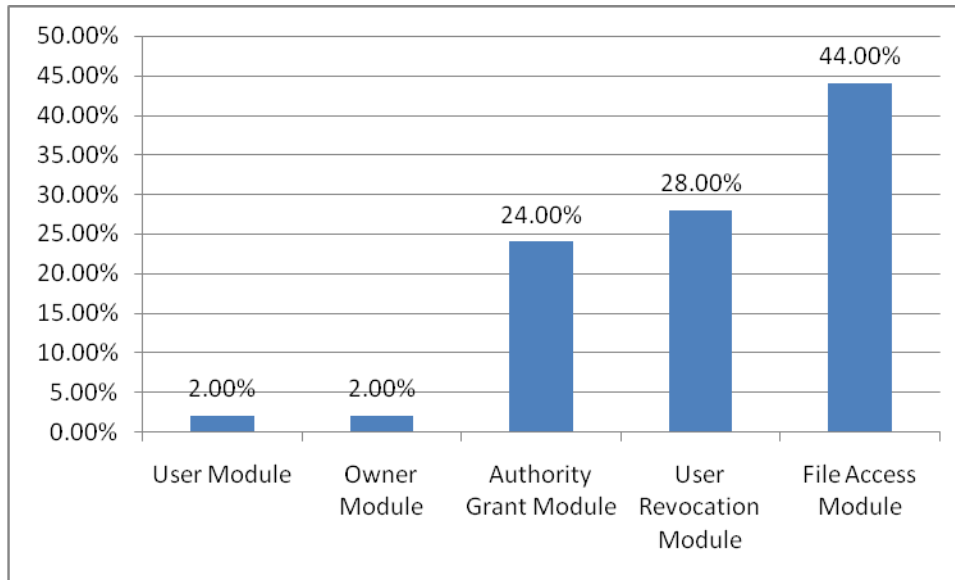
From the above table and chart, it is interpreted that 48.00% of the respondents are dissatisfied, 26.00% of the respondents are neutral, 22.00% of the respondents are satisfied and 2.00% of the respondents are highly satisfied and highly dissatisfied based on the satisfaction with the skills of the trainers of the data outsourcing process.

TABLE NO 4.1.17

Respondents classification based on the most important module of data outsourcing

SI.No	Description	Frequency	Percentage
1	User Module	1	2.00%
2	Owner Module	1	2.00%
3	Authority Grant Module	12	24.00%
4	User Revocation Module	14	28.00%
5	File Access Module	22	44.00%
	Total	50	100%

CHART NO 4.1.17



Interpretation

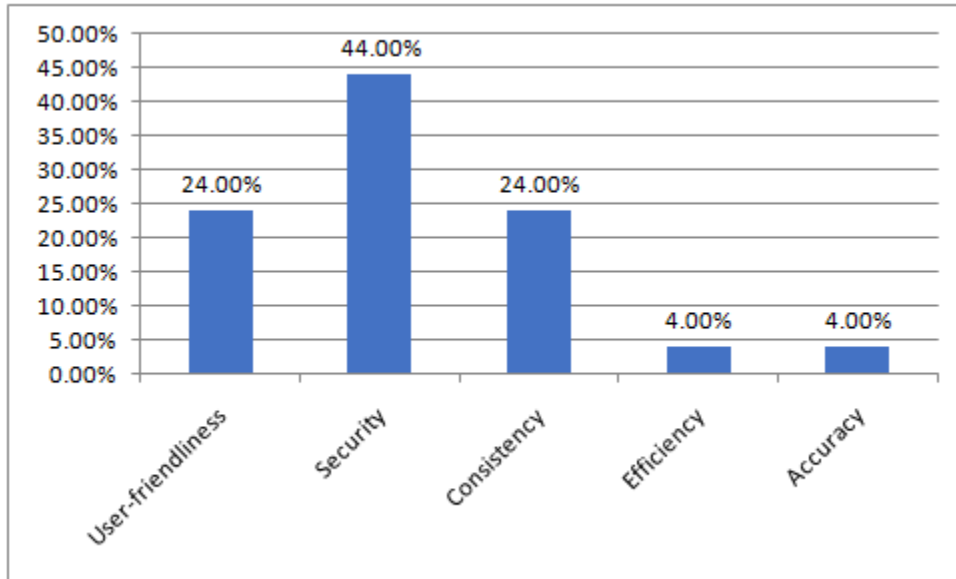
From the above table and chart, it is interpreted that 44.00% of the respondents as file access module, 28.00% of the respondents as user revocation module, 24.00% of the respondents as authority grant module and 2.00% of the respondents as user module and owner module based on the most important module of data outsourcing.

TABLE NO 4.1.18

Respondents classification based on the most important feature of the data outsourcing

SI.No	Description	Frequency	Percentage
1	User-friendliness	12	24.00%
2	Security	22	44.00%
3	Consistency	12	24.00%
4	Efficiency	2	4.00%
5	Accuracy	2	4.00%
	Total	50	100%

CHART NO 4.1.18



Interpretation

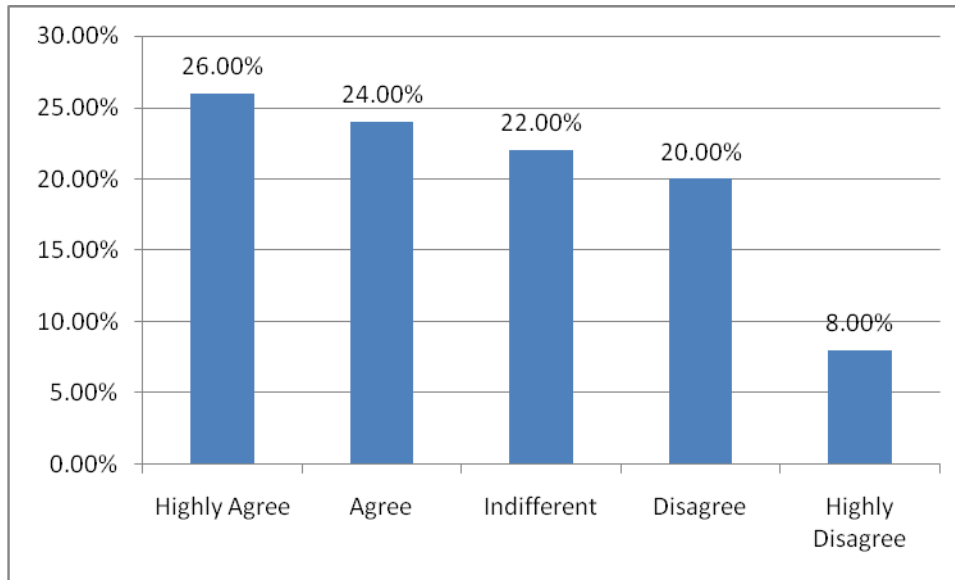
From the above table and chart, it is interpreted that 44.00% of the respondents as security and 24.00% of the respondents have interpreted as user-friendliness, consistency and 4.00% of the respondents as efficiency, accuracy based on the most important feature of the data outsourcing.

TABLE NO 4.1.19

Respondents classification based on the algorithm used is best in the organization

SI.No	Description	Frequency	Percentage
1	Highly Agree	13	26.00%
2	Agree	12	24.00%
3	Indifferent	11	22.00%
4	Disagree	10	20.00%
5	Highly Disagree	4	8.00%
	Total	50	100%

CHART NO 4.1.19



Interpretation

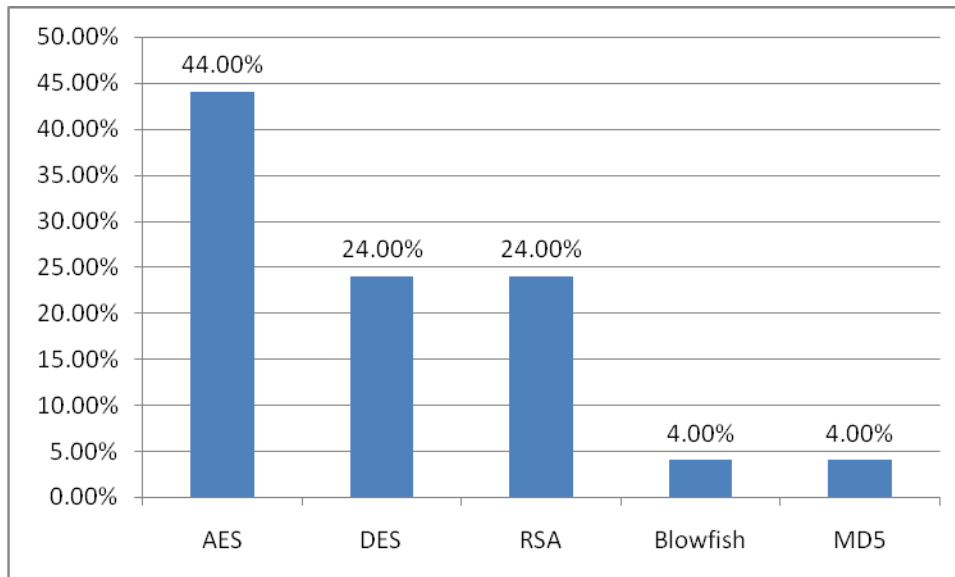
From the above table and chart, it is interpreted that 26.00% of the respondents highly agree, 24.00% of the respondents agree, 22.00% of the respondents are indifferent, 20.00% of the respondents disagree and 8.00% of the respondents highly disagree based on the algorithm used is best in the organization.

TABLE NO 4.1.20

Respondents classification based on the security algorithm can be suggested to the organization for the better usage

Sl.No	Description	Frequency	Percentage
1	AES	22	44.00%
2	DES	12	24.00%
3	RSA	12	24.00%
4	Blowfish	2	4.00%
5	MD5	2	4.00%
	Total	50	100%

CHART NO 4.1.20



Interpretation

From the above table and chart, it is interpreted that 44.00% of the respondents have suggested AES algorithm, 24.00% of the respondents has suggested DES and RSA as the security algorithm can be suggested to the organization for the better usage.

4.2 SWOT ANALYSIS

The SWOT Analysis represents the strength, weakness, opportunities and threats that are involved in the secured process of data outsourcing.

STRENGTH

The strengths that are involved in the process of secured data outsourcing are:

- Highest Level of Security
- Scalability
- Easy Access

WEAKNESS

The weakness that are involved in the process of secured data outsourcing are:

- Requires Technical Knowledge
- Cost expensive system

OPPORTUNITIES

The opportunities that are involved in the process of secured data outsourcing are:

- Day by Day Increasing Number of Users
- Increasing number of data
- Increasing number of challenges

THREATS

The threats that are involved in the process of secured data outsourcing are:

- Hackers
- Data Loss

CHAPTER – 5

SUMMARY, CONCLUSION AND SUGGESTIONS

5.1 SUMMARY OF FINDINGS

- From the analysis, it is found that 48.00% of the respondents are between 31-35 yrs, 26.00% of the respondents are below 30 Yrs, 16.00% of the respondents are between 36-40 yrs and 8.00% of the respondents are above 46 yrs and above.
- From the analysis, it is found that 52.00% of the respondents are females and 48.00% of the respondents are male respectively.
- From the analysis, it is found that 38.00% of the respondents are undergraduates, 30.00% of the respondents are engineering graduates, 22.00% of the respondents are diploma holders and 6.00% of the respondents are post graduates.
- From the analysis, it is found that 28.00% of the respondents are below 3 years, 26.00% of the respondents are between 4 to 6 yrs, 24.00% of the respondents are between 6 to 8 yrs and 16.00% of the respondents are above 10 yrs of experience in information security.
- From the analysis, it is found that 28.00% of the respondents are getting monthly income as 10001-20000, 26.00% of the respondents are getting monthly income as 20001-30000, 24.00% of the respondents are getting monthly income as 30001-50000, 10.00% of the respondents as 50001-60000 and 12.00% of the respondents are getting monthly income less than above 10000.
- From the analysis, it is found that 44.00% of the respondents as AES, 24.00% of the respondents as DES, 22.00% of the respondents as RSA, 6.00% of the

respondents as Blowfish and 4.00% of the respondents as MD5 as the algorithm that will be the best suggested.

- From the analysis, it is found that 40.00% of the respondents as more than 3 years, 28.00% of the respondents as 2-3years, 26.00% of the respondents as 1-2years, 4.00% of the respondents as less than 1 year and 2.00% of the respondents are not aware based on when the server will be usually upgraded in the company.
- From the analysis, it is found that 30.00% of the respondents as 6 months – 1 year, 28.00% of the respondents as less than 6months and 1-2years, 12.00% of the respondents as 2-3 years and 2.00% of the respondents as more than 3 years based on starting to upload the confidential data.
- From the analysis, it is found that 50.00% of the respondents strongly agree, 30.00% of the respondent agree, 10.00% of the respondent disagree, 6.00% of the respondent strongly disagree and 4.00% of the respondent are neutral based on ensuring the security of the data in the organization.
- From the analysis, it is found that 30.00% of the respondents disagree, 26.00% of the respondents strongly disagree, 20.00% of the respondents strongly agree, 16.00% of the respondents agree and 8.00% of the respondents are neutral based on expecting the management to provide security of data.
- From the analysis, it is found that 44.00% of the respondents as AES, 24.00% of the respondents as RSA, 22.00% of the respondents as DES, 6.00% of the respondents as MD5 and 4.00% of the respondents as blowfish based on the algorithm that is presently used for secure data outsourcing.
- From the analysis, it is found that 40.00% of the respondents are satisfied, 28.00% of the respondents are dissatisfied, 26.00% of the respondents are neutral, 4.00% of the respondents are highly satisfied and 2.00% of the respondents are highly

dissatisfied based on the satisfaction with the performance of the data outsourcing.

- From the analysis, it is found that 38.00% of the respondents are highly satisfied, 28.00% of the respondents are satisfied, 26.00% of the respondents are neutral and 4.00% of the respondents are dissatisfied and highly dissatisfied based on the satisfaction with the depth of the working of the security algorithm.
- From the analysis, it is found that 38.00% of the respondents are satisfied, 28.00% of the respondents are highly satisfied, 26.00% of the respondents are neutral and 4.00% of the respondents are dissatisfied and highly dissatisfied based on the satisfaction with the training methods of the working of the algorithm.
- From the analysis, it is found that 36.00% of the respondents are satisfied, 28.00% of the respondents are neutral, 26.00% of the respondents are dissatisfied, 6.00% of the respondents are highly satisfied and 4.00% of the respondents are highly dissatisfied based on the satisfaction with the quality of training of the data outsourcing process.
- From the analysis, it is found that 48.00% of the respondents are dissatisfied, 26.00% of the respondents are neutral, 22.00% of the respondents are satisfied and 2.00% of the respondents are highly satisfied and highly dissatisfied based on the satisfaction with the skills of the trainers of the data outsourcing process.
- From the analysis, it is found that 44.00% of the respondents as file access module, 28.00% of the respondents as user revocation module, 24.00% of the respondents as authority grant module and 2.00% of the respondents as user module and owner module based on the most important module of data outsourcing.

- From the analysis, it is found that 44.00% of the respondents as security and 24.00% of the respondents have interpreted as user-friendliness, consistency and 4.00% of the respondents as efficiency, accuracy based on the most important feature of the data outsourcing.
- From the analysis, it is found that 26.00% of the respondents highly agree, 24.00% of the respondents agree, 22.00% of the respondents are indifferent, 20.00% of the respondents disagree and 8.00% of the respondents highly disagree based on the algorithm used is best in the organization.
- From the analysis, it is found that 44.00% of the respondents have suggested AES algorithm, 24.00% of the respondents has suggested DES and RSA as the security algorithm can be suggested to the organization for the better usage.

5.2 CONCLUSION

This project presented a security based data outsourcing that is significantly faster than existing approaches. This system has provided flexibility of search to the users and at the same time this project provided a secure and trustable approach for the data owner as well. Cryptography is implemented in order to provide security and privacy for the data. AES algorithm is used to ensure the security of the data. It is beneficial both to the data owner as well as the user. The result shows that this approach is better in terms of storage, communication and outsourcing of data.

5.3 SUGGESTIONS AND RECOMMENDATIONS

From the above findings, the below suggestions and recommendations are provided.

- Understand that the company can be a target to hackers.
- Keep software up to date
- Avoid Phishing scams - beware of suspicious emails and phone calls
- Practice good password management
- Be careful in clicking links
- Never leave devices unattended
- Protect sensitive data
- Use mobile devices safely
- Install anti-virus protection
- Back up your data

REFERENCES

- [1] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in In proceedings of Eurocrypt, 2004, pp. 506–522.
- [2] B. Waters, D. Balfanz, G. Durfee, and D. K. Smetters, “Building an encrypted and searchable audit log,” in Network and Distributed System Security Symposium, 2004.
- [3] M. Ding, F. Gao, Z. Jin, and H. Zhang, “An efficient public key encryption with conjunctive keyword search scheme based on pairings,” in IEEE International Conference on Network Infrastructure and Digital Content, 2012, pp. 526–530.
- [4] F. Kerschbaum, “Secure conjunctive keyword searches for unstructured text,” in International Conference on Network and System Security, 2011, pp. 285–289.
- [5] C. Hu and P. Liu, “Public key encryption with ranked multi keyword search,” in International Conference on Intelligent Networking and Collaborative Systems, 2013, pp. 109–114.
- [6] Z. Fu, X. Sun, N. Linge, and L. Zhou, “Achieving effective cloud search services: multi-keyword ranked search over encrypted cloud data supporting synonym query,” IEEE Transactions on Consumer Electronics, vol. 60, pp. 164–172, 2014.
- [7] C. L. A. Clarke, G. V. Cormack, and E. A. Tudhope, “Relevance ranking for one to three term queries,” Information Processing and Management: an International Journal, vol. 36, no. 2, pp. 291–311, Jan. 2000.
- [8] H. Tuo and M. Wenping, “An effective fuzzy keyword search scheme in cloud computing,” in International Conference on Intelligent Networking and Collaborative Systems, 2013, pp. 786–789.
- [9] M. Zheng and H. Zhou, “An efficient attack on a fuzzy keyword search scheme over encrypted data,” in International Conference on High Performance Computing and Communications and Embedded and Ubiquitous Computing, 2013, pp. 1647–1651.
- [10] S. Zittrower and C. C. Zou, “Encrypted phrase searching in the cloud,” in IEEE Global Communications Conference, 2012, pp. 764–770.

APPENDIX

QUESTIONNAIRE

1) Age

- a) Below 30 Yrs
- b) 31 Yrs – 35 Yrs
- c) 36 Yrs – 40 Yrs
- d) 41 Yrs -45 Yrs
- e) 46 Yrs and above

2) Gender

- a) Male b) Female

3) Educational Qualification

- a) Diploma Holder
- b) Engineering Graduate
- c) Under Graduation
- d) Post Graduation
- e) Others

4) Experience in Information Security

- a) Below 3 Yrs
- b) 4 to 6 Yrs
- c) 6 to 8 Yrs
- d) 8 to 10 Yrs
- e) Above 10 Yrs

5) Monthly Income (in Rs)

- a) Less than 10000
- b) 10001 $\pm$ 20000
- c) 20001 $\pm$ 30000
- d) 30001 $\pm$ 50000
- e) More than 50000

6) Which algorithm will be the best suggested?

- a) AES
- b) DES
- c) RSA
- d) Blowfish
- e) MD5

7) When will be the data usually uploaded in the company?

- a) Not aware
- b) Less than 1 year
- c) 1-2 years
- d) 2-3 years
- e) More than 3 years

8) When did you start to upload the data?

- a) Less than 6 months ago
- b) 6months-1year
- c) 1-2years
- d) 2-3years
- e) More than 3years

9) The organization ensures the security of the data

- a) Strongly Agree
- b) Agree
- c) Neutral
- d) Disagree
- e) Strongly Disagree

10) You expect the management to provide the security of data

- a) Strongly Agree
- b) Agree
- c) Neutral
- d) Disagree
- e) Strongly Disagree

11) Which algorithm is presently used for secure data outsourcing?

- a) AES
- b) DES
- c) RSA
- d) Blowfish
- e) MD5

12) You are satisfied with the performance of the data outsourcing

- | | | |
|---------------------|------------------------|------------|
| a) Highly Satisfied | b) Satisfied | c) Neutral |
| d) Dissatisfied | e) Highly Dissatisfied | |

13) You are satisfied with the depth of the working of the security algorithm

- | | | |
|---------------------|------------------------|------------|
| a) Highly Satisfied | b) Satisfied | c) Neutral |
| d) Dissatisfied | e) Highly Dissatisfied | |

14) You are satisfied with the training methods of the working of the algorithm

- | | | |
|---------------------|------------------------|------------|
| a) Highly Satisfied | b) Satisfied | c) Neutral |
| d) Dissatisfied | e) Highly Dissatisfied | |

15) You are satisfied with the quality of training of the data outsourcing process

- | | | |
|---------------------|------------------------|------------|
| a) Highly Satisfied | b) Satisfied | c) Neutral |
| d) Dissatisfied | e) Highly Dissatisfied | |

16) You are satisfied with the skills of the trainers of the data outsourcing process

- | | | |
|---------------------|------------------------|------------|
| a) Highly Satisfied | b) Satisfied | c) Neutral |
| d) Dissatisfied | e) Highly Dissatisfied | |

17) Which is the most important module of data outsourcing?

- a) User Module
- b) Owner Module
- c) Authority Grant Module
- d) User Revocation Module
- e) File Access Module

18) Which is the most important feature of the data outsourcing?

- a) User-friendliness
- b) Security
- c) Consistency
- d) Efficiency
- e) Accuracy

19) The algorithm used is best in the organization

- a) Strongly Agree b) Agree c) Neutral
- d) Disagree e) Strongly Disagree

20) Which security algorithm can be suggested to the organization for the better usage?

- a) AES
- b) DES
- c) RSA
- d) Blowfish
- e) MD5