# Enabling the search using fine Grained multi keyword search algorithm over encrypted cloud data

B.Muthulakshmi[#1], R.Sathyapriya[#2], J.Janani[#3], K.Abilasha[#4]
[#1]*Assistant Professor, A.V.C.College of Engineering*
[#2-4]*UG Students, A.V.C.College of Engineering*

***Abstract—*** *Using cloud computing, individuals can store their data on remote servers and allow data access to public users through the cloud servers. As the outsourced data are likely to contain sensitive privacy information, they are typically encrypted before uploaded to the cloud. In this paper, we address this issue by developing the fine-grained multi-keyword search schemes over encrypted cloud data. Our original contributions are three-fold. First, we are design the cloud framework for multi data owner. Second, we construct index vectors based on fuzzy search approach then we use Efficient and semantic based keyword search. Third, Decrypt the data by using Verification scheme for outsourced cloud storage.*

***Keywords—*** *Index table, Searchable encryption, multi-keyword, fine-grained, cloud computing.*

## I. INTRODUCTION

The cloud computing treats computing as a utility and leases out the computing and storage capacities to the public individuals . In this framework, the individual can remotely store her data on the cloud server, namely data outsourcing, and then make the cloud data open for public access through the cloud server. This represents a more scalable, low-cost and stable way for public data access because of the scalability and high efficiency of cloud servers, and therefore is favorable to small enterprises and the outsourced data may contain sensitive privacy information. It is often necessary to encrypt the private data before transmitting the data to the cloud servers. The data encryption, however, would significantly lower the usability of data due to the difficulty of searching over the encrypted data .. Simply encrypting the data may still cause other security concerns. For instance, Google Search uses SSL (Secure Sockets Layer) to encrypt the connection between search user and Google server when private data, such as documents and emails, appear in the search results However, if the search user clicks into another website from the search results page, that website may be able to identify the search terms that the user has used. On addressing this issues, the searchable encryption has been recently developed as a fundamental approach to enable searching over encrypted cloud data, which proceeds the following operations. First, the data owner needs to generate several keywords according to the outsourced data. These keywords are then encrypted and stored at the cloud server. When a search user needs to access the outsourced data, it can select some relevant keywords and send the cipher text of the selected keywords to the cloud server. The cloud server then uses the cipher text to match the outsourced encrypted keywords, and lastly returns the matching results to the search user. We introduce the relevance scores and the preference factors of keywords for searchable encryption. The relevance scores of keywords can enable more precise returned results, and the preference factors of keywords represent the importance of keywords in the search keyword set specified by search users and correspondingly enable personalized search to cater to specific user preferences. This paper a Multi data owner scheme can be implement in cloud storage. That is used for Semantic search for retrieving documents using fuzzy search and implement ECC based algorithm to encrypt the cloud data. The Authentication scheme can also be implemented to separate public key into two parts at the time of data downloading. Using these schemes , we can retrieve large amount of relevant documents based on user searchable keyword and anonymous access can be blocked. Mapping can be done in encrypted cloud storage and relationship between multi- keywords can be analyzed easily.

## II. RELATED WORK

**C. Wang** et.al [5] proposed a model solve the challenging problem of privacy-preserving multi-

keyword ranked search over encrypted data in cloud computing (MRSE). We establish a set of strict privacy requirements for such a secure cloud data utilization system. Among various multi-keyword semantics, we choose the efficient similarity measure of ―coordinate matching,‖ i.e., as many matches as possible, to capture the relevance of data documents to the search query. We further use ―inner product similarity‖ to quantitatively evaluate such similarity measure.

**M. M. Mahmoud** et.al [2] proposed a model a hotspot phenomenon that causes an obvious inconsistency in the network traffic pattern due to the large volume of packets originating from a small area. Second, we develop a realistic adversary model, assuming that the adversary can monitor the network traffic in multiple areas, rather than the entire network or only one area. Using this model, we introduce a novel attack called Hotspot-Locating where the adversary uses traffic analysis techniques to locate hotspots.

**H. Liang** et.al [3] proposed a service decision making system for inter domain service transfer to balance the computation loads among multiple cloud domains. Our system focuses on maximizing the rewards for both the cloud system and the users by minimizing the number of service rejections that degrade the user satisfaction level significantly. To this end, we formulate the service request decision making process as a semi-Markov decision process[SMDP].

**T. Jung** et.al [4] proposed a model external aggregator or multiple parties can learn some algebraic statistics (e.g., sum, product) over participants' privately owned data while preserving the data privacy. Assume all channels are subject to eavesdropping attacks, and all the communications Throughout the aggregation are open to others. Using several protocols that successfully guarantee data privacy under this weak assumption while limiting both the communication and computation complexity of each participant to a small constant.

**Q. Shen** et.al [1] proposed a model e-health monitoring system with minimum service delay and privacy preservation by exploiting geo-distributed clouds. In the system, the resource allocation scheme enables the distributed cloud servers to cooperatively assign the servers to the requested users under the load balance condition. Thus, the service delay for users is minimized. In addition, a traffic-shaping algorithm is proposed. The traffic-shaping algorithm converts the user health data traffic to the non health data traffic such that the capability of traffic analysis attacks is largely reduced.

## III. MODULES

Users Initialization
File upload
Index table construction
Search based on keywords
Verificable outsourced decryption

System Architecture

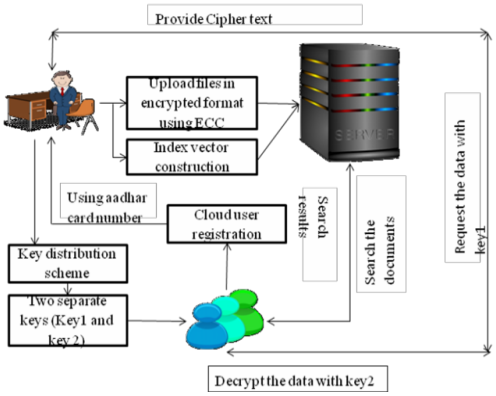


**Fig.1: System Architecture 1.Users Initialization:**

In this module we initialize the cloud with following members such as
Cloud owner
Cloud server
User

Cloud owner
The data owner outsources her data to the cloud for convenient and reliable data access to the corresponding search users. To protect the data privacy, the data owner encrypts the original data through symmetric encryption. To improve the search efficiency, the data owner generates some keywords for each outsourced document. The corresponding index is then created according to the keywords and a secret key.

Cloud server
The cloud server is an intermediate entity which stores the encrypted documents and corresponding indexes that are received from the data owner, and provides data access and search services to search users. When a search user sends a keyword trapdoor to the cloud server, it would return a collection of matching documents based on certain operations.

Cloud user
A search user queries the outsource documents from the cloud server with following three steps.

First, the search user receives both the secret key and symmetric key from the data owner. Second, according to the search keywords, the search user uses the secret key to generate trapdoor and sends it to the cloud server. Lastly, user receives the matching document collection from the cloud server and decrypts them with the symmetric key.

File upload:

We can upload the files into cloud system. then, files can be encrypted using ECC algorithm.ECC base keys can be generated and stored in cloud system.

Index table construction:

We can construct index table based on keywords in the priority vise. Keywords are trained based on user search. Calculate the term frequency and term similarity. Rank the keywords and also analyze misspelling words.

INDEX TABLE CONSTRUCTION:

I Table Index Table construction

| OWNER NAME | FILE NAME | POSSIBLE KEYWORDS | RANK |
|---|---|---|---|
| Sathu | Compute r network | Protocol | 3 |
| Janani | Database management | Data | 1 |
| Abilasha | os | unix | 2 |

II Table RANKING KEYWORDS:

| KEYWORDS | PRIORITY |
|---|---|
| DATA | 10 |
| SERVER | 8 |
| PROTOCOL | 5 |
| INFORMATION | 4 |

## IV. SEARCHING BASED ON KEYWORDS:

Cloud users can register into the cloud owner based on their aadhar card number. Search in cloud based on keyword or multi-keyword or more than one word. Words are searched in index table and to provide search results with data owner details

## V. VERIFICABLE OUTSOURCED DECRYPTION:

This module can be used to provide two keys to users after registration. Based on two keys, user can request the data with key1.After that, user get cipher text from data owner. Decrypt the data based on another key.

A. ECC ALGORITHM DEFINITION:

Elliptical curve cryptography (ECC) is a public key encryption technique based on *elliptic curve theory* that can be used to create faster, smaller, and more efficient cryptographic keys. ECC generates keys through the properties of the elliptic curve equation instead of the traditional method of generation as the product of very large prime numbers. The technology can be used in conjunction with most public key encryption methods, such as RSA, and Diffie-Hellman. According to some researchers, ECC can yield a level of security with a 164-bit key that other systems require a 1,024-bit key to achieve. Because ECC helps to establish equivalent security with lower computing power and battery resource usage, it is becoming widely used for mobile applications.

ALGORITHM:

Algorithm 1: Elliptic Curve Encryption

Input: Parameters from the elliptic curve domain (p, E, P, n), Public Key Q,
Raw Text m Output: Encrypted text (C1, C2) begin
Represent the message m as a point M in E(Fp)
Select k ∈R[1,n−1].
Calculate C1 = kP
Calculate C2 = M + kQ.
Return (C1, C2) end.

Algorithm 2: Elliptic Curve Decryption

Input: Parameters from the elliptic curve domain (p, E, P, n), Private key d, Encrypted text (C1, C2)

Output: Raw Text m begin
Calculate M = C2- dC1 and extract m from M.
Return (m) end.

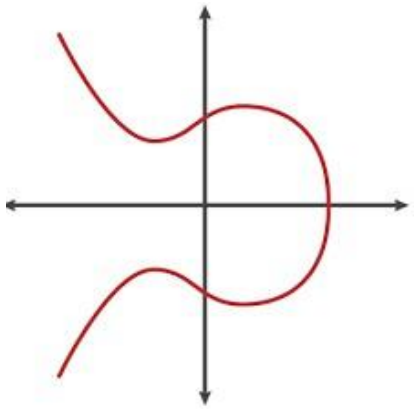

**Fig.2: Elliptic curve cryptography**

## VI. ADVANTAGES OF ELLIPTIC CURVE:

Formation of Curve .Shorter keys are as strong as long key for RSA.Low on CPU consumption.Low on memory usage.

## VII. CONCLUSION

we have investigated on the Efficient fine grained access control with semantic keyword search on encrypted cloud storage. We are provide access control mechanisms for retrieving data from cloud with verifiable outsourced decryption in encrypted cloud system. For the future work, we intend to further extend the proposal to consider the extensibility of the file set and the multi-user cloud environments. Towards this direction, we have made some preliminary results on the extensibility and the multi-user cloud environments. Another interesting topic is to develop the highly scalable searchable encryption to enable efficient search on large practical databases.

## REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, ―Privacy- preserving multi-keyword ranked search over encrypted cloud data,‖ IEEE Trans. Parallel Distrib. Syst., vol. 25, no. 1, pp. 222–233, Jan. 2014.

[2] M. M. Mahmoud and X. Shen, ―A cloud-based scheme for protecting source-location privacy against hotspot-locating attack in wireless sensor networks,‖ IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 10, pp. 1805–1818, Oct. 2012.

[3] Q. Shen, X. Liang, X. Shen, X. Lin, and H. Luo, ―Exploiting geodistributed clouds for e-health monitoring system with minimum service delay and privacy preservation,‖ IEEE J. Biomed. Health Inform., vol. 18, no. 2, pp. 430–439, Mar. 2014.

[4]T. Jung, X. Mao, X. Li, S.-J. Tang, W. Gong, and L. Zhang, ―Privacy-preserving data aggregation without secure channel:Multivariate polynomial evaluation,‖ in Proc. IEEE INFOCOM, 2013, pp. 2634–2642.

[5] H. Liang, L. X. Cai, D. Huang, X. Shen, and D. Peng, ―An smdpbased
service model for interdomain resource allocation in mobile cloud networks,‖ IEEE Trans. Veh. Technol., vol. 61, no. 5, pp. 2222–2232, Jun. 2012.