FILE SEARCH BASED ON RANKING SCHEME OVER ENCRYPTED CLOUD

Ms.M.Malathi[1], *S.Sathya Priya* [2], *S.Surekha* [3]
Assistant Professor, Department of Information Technology, Sri Krishna College of Technology, Kovaipudur, Coimbatore, India[1]

U.G Scholar, Department of Information Technology, Sri Krishna College of Technology, Kovaipudur, Coimbatore, India[2,3]

ABSTRACT –

A Secure multi keyword File Searching based on Ranking Scheme over Encrypted Cloud is about the data owners are motivated to encrypt their data and it is hosted to the cloud servers for the better usability and to maintain the data at the lowest cost. Therefore, sensitive data should be encrypted for privacy requirements. This reduces the hacking of files from the unauthorized user access. In this paper, we present a secure multi keyword file searching based on ranking scheme over encrypted cloud data. This simultaneously supports dynamic operations like deletion or insertion of files. We have proposed a technique called "Fuzzy keyword" algorithm to provide efficient multi-keyword file search. The secure SHA algorithm is utilized to encrypt the files that is being loaded by the data owner. Due to the encryption of files, the scheme achieves linear activities such as searching a particular file and it also deals with major operations such as deletion or insertion of documents effectively.

Index Terms- Searchable encryption, multi-keyword file search, cloud computing.

I INTRODUCTION

Cloud computing has considered as a new model of IT infrastructure, which can be used to organize a huge resource of computing, storage. It enables the users to enjoy enormous, convenient and on-demand network access to a pool of computing resources with qualified efficiency and with minimised economic rate. Both the individuals and enterprises are motivated to outsource their data to the cloud, instead of purchasing the software or hardware and installing on to their computer.

Due to the various rewards of cloud services, encrypted sensitive information or data to the remote servers brings as many privacy concerns as possible. The cloud service providers(CSP) may access the user's sensitive information without their authorization. A general approach to protect the data from third party users is to encrypt the data before outsourcing. This might cause a huge cost in terms of data encryption as there is a need for installing a software or hardware. Searching of files in a encrypted coud data in not applicable. Downloading of documents without decrypting the files is quite difficult.

In order to overcome the above problem, the developers designed some general-purpose solutions for solving the issue of downloading the documents in a much easier way. This practical special solution, such as searchable encryption scheme(SE) have made contributions like efficiency, availabilty, security and functionality. Searchable encryption scheme means that allows the clients to store the encrypted files onto the cloud and execute the keyword file search over the ciphertext domain. So far, abundant works has been proposed to obtain various search functionalities. Among them, multi keyword search takes up more attention for its availability. It also includes dynamic operations such as insertion, updation and deletion of documents.

This paper specifically includes the construction of encrpyted index using "SHA algorithm" and also "Fuzzy keyword algorithm" is used for performing secure multi-keyword file search. The file gets encrypted as soon as the owner uploads the data onto the cloud. When a user login to the cloud to download a specific file, the user has to mention the file name and a key will be generated to the registered mail id. The user has to enter the key in order to download the respective file. Our contributions are summarized as follows:

1. The SHA algorithm is used to encrypt the files that is loaded on the cloud server.

Fuzzy keyword algorithm is used to produce the needed file that is searched in a ranked based scheme.

II RELATED WORK

Searchable encryption scheme enables storage of encrypted data on the cloud and helps to execute a keyword based file search over the ciphertext domain. There are so many encryption scheme that has been developed by different publishers.

1. **"Secure and Private Cloud-Based Control Using Semi- Homomorphic Encryption"**

The work was, in part, supported by a McKenzie Fellowship, ARC grant and Defence Science and Technology Group through the Research Agreement My IP:6288.

In this paper, the parameters of the Paillier encryption technique are determined. This improves the stability of the closed-loop performance.

- The semi-homomorphic or homomorphic encryption schemes are utilized when using third-party cloud services.

2. **"Cryptographic Cloud Storage with security and data sharing for the Multi access network"**

The major aim of this technique a secure multi-owner information sharing scheme. It tells that the user will share the inforamtion within the cluster by the trustworthy cloud in the world organisation.

3. **"Fully homomorphic encryption scheme"**

This scheme is used to perform few functions over encrypted data without the help of decryption key – i.e., given encryptions E(m1), . . . , E(mt) of m1, . . . , mt , one can efficiently compute a compact ciphertext that encrypts f(m1, . . . , mt) for any efficiently computable function. it enables private queries to a search engine.

4. **"Searchable Symmetric Encryption with Improved Efficient Constructions and Definitions"**

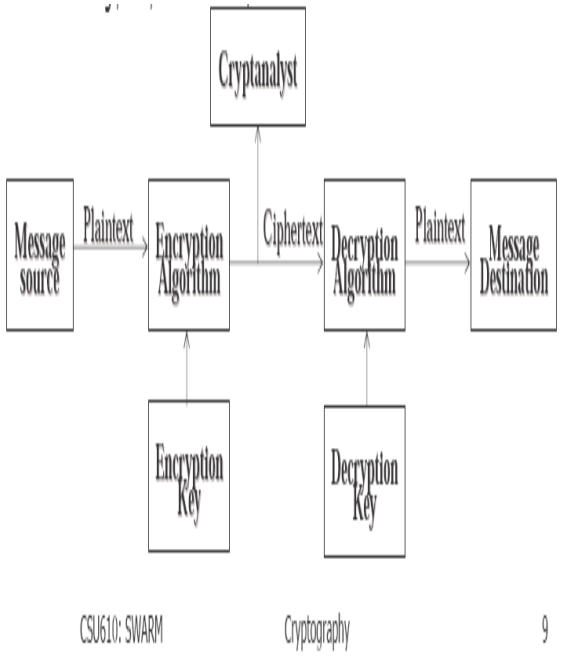
Searchable symmetric encryption (SSE) involves in a process that allows a party to outsource the storage of the data to another party in a private manner, while maintaining the ability to selectively search over it.

5. **"Privacy Preserving Keyword Searches on Remote Encrypted Data"**

The scheme is efficient in the sense that no publickey cryptosystem is involved. A user can store his/her files in an encrypted form on a remote file server. Later the user can efficiently retrieve some of the encrypted files containing (or indexed by) specific keywords, keeping the keywords themselves secret.

The existing schemes on keyword-based file retrieval are widely used on the plain text data and cannot be directly applied on the encrypted data. All these file search schemes retrieve search results based on the ranking of keywords, which cannot provide acceptable result. Therefore, the files should be encrypted before outsourcing for privacy needs.

In this proposed system, we use two algorithm namely SHA algorithm for encrypting the files and Fuzzy algorithm for searching a file using multi keyword based on ranking scheme over an encrypted cloud. The files can be downloaded by the data user from the cloud server using a auto generated keyword that has been send to the registered mail id of the data user. This becomes more efficient in searching a file.



III PROBLEM FORMULATION

SYSTEM ARCHITECTURE

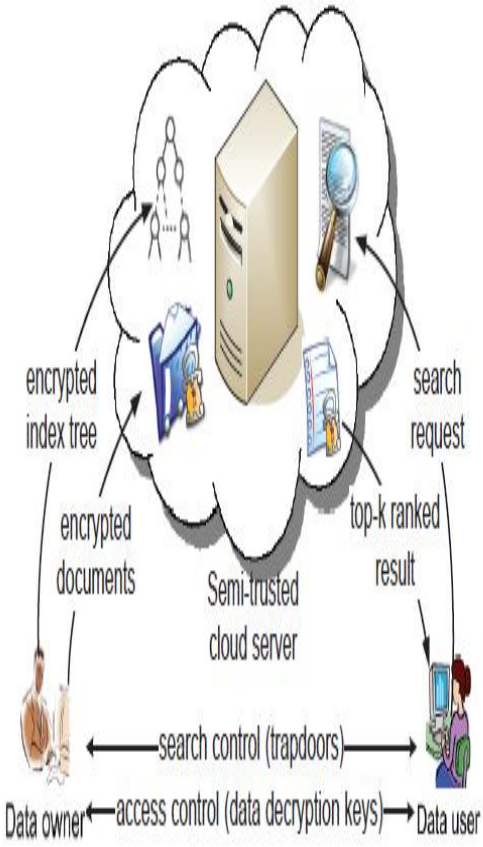
The system architecture encampasses three different roles: data owner, data user and cloud server.

Data owner contains the collection of documents that the owner wants to outsource to the server in encrypted form. In addition he also concentrates in securing his capability to search on them for effective utilization. This paper priorities the process in the following manner**.**

1. Builds a secure searchable index from file collection.
2. Generates an encrypted file collection.
3. Data owner outsources the encrypted files and secure index to the cloud server.
4. Distributes the key information to the authorised data users.

Data users are the authorised data users who can access the encrypted document by using the specific keyword. For this to be done, the data user has to register the nameof the document in the cloud server. By registering, the document gets uploaded so that each time the data user can login the cloud server and enter the file name. This makes the server to send a keyword to the data users mail id. By using this keyword ,the user can access the encrypted document and then he can decrypt.

Cloud server is a online storage server that is used to store a huge amount of files or data. The datas gets encrypted when it is uploaded on the server. This is used to overcome the security purpose from the third party authorities. This server is the major base used to store the encrypted document collection. The cloud server can be used from anywhere, at anytime. This reduces the huge cost as there is no need of downloading or of installing any hardware or software.



DESIGN GOALS

The encrypted cloud has to provide the accurate,secure,efficient and dynamic multi keyword file search based on ranking scheme. Our system undergoes the following design goals:

Dynamic: The scheme has been designed to provide not only the multi-keyword file search and also accurate result ranking, and it also involve the dynamic updation in the documents.

Search efficiency: The concept aims at providing a very basic server for exploring the encrypted documents using a unique index and also by employing efficient search algorithm.

Privacy preserving: The scheme is specifically designed to prevent the cloud server from accessing the additional information about the document as well as the index and the query. Thus it provides a secure environment. The privacy requirements are performed through index and query confidentiality and the keyword privacy.

IV PROPOSED SCHEME

File Search based on Ranking Scheme over Encrypted Cloud is constructed by encrypting the files using a SHA algorithm and also proposed a "Fuzzy Search" algorithm to provide an efficient file search. The proposed scheme can achieve simultaneous operations like searching a file and it also deals with the deletion and insertion of documents easily. Experiments have been conducted to establish the efficiency of the scheme.

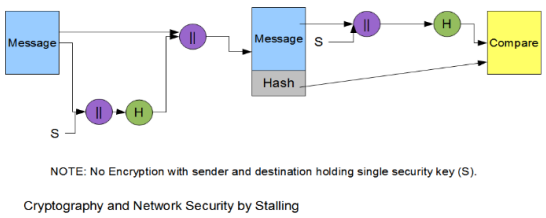
- Abundant works has been proposed to achieve various search functionalities.
- At present, few dynamic methods have been proposed to support the insertion and deletion operations in a document.

This paper proposes a secure multi keyword file search scheme over the encrypted cloud, which simultaneously supports secure dynamic operations like insertion and deletion of content on the document collection.

ENCRYPTION ALGORITHM

SHA- Secure Hash Algorithm

In 1993, the specification of the algorithm has been published as the *Secure Hash Standard*, FIPS PUB 180, by US government standards agency NIST (National Institute of Standards and Technology). This version is known as "SHA-0". SHA-1 differs from SHA-0 from its compression function only by a single bitwise rotation. Weaknesses have subsequently has been reported in both SHA-0 and SHA-1. SHA-1 has a feature to provide greater resistance against the attacks that is faced by the documents.
SHA-1 (as well as SHA-0) is based on principles that is similar to those used by Professor Ronald L. It produces a 160-bit digest from a message.



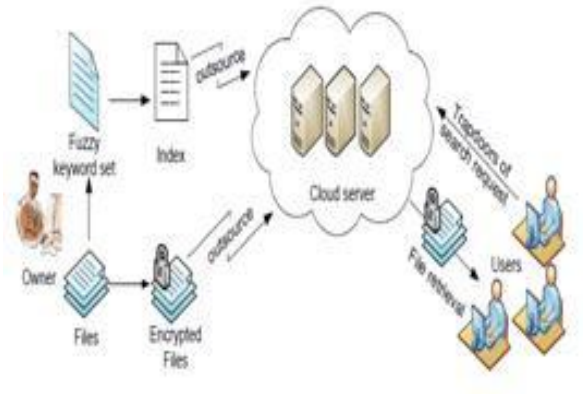
NOTE: No Encryption with sender and destination holding single security key (S).

Cryptography and Network Security by Stalling

SHA is used for the encryption purpose and it is the most widely used algorithm for searchable encryption scheme. Fuzzy keyword algorithm is used for retriving the files based on the ranking scheme.

RANKING PROCESS

The search process is implemented using a ranked scheme and this is done using a algorithm called "Fuzzy keyword algorithm". A **fuzzy concept** is a concept that varies according to the condition or the context in a document. This means the concept is lacking in a fixed manner. It is vague in few concepts. It provides a precise meaning, without being unclear or meaningless altogether. On further elaboration and specification it has a definite meaning, which can be more precised.



V SYSTEM DESIGN AND IMPLEMENTATION

The proposed model is for handling a secure multi keyword file search based on ranking scheme over an encrpyted cloud data collection. SHA algorithm is used for encrypting the files that is uploaded on the cloud and Fuzzy algorithm is used for searching a file based on ranking scheme. This simultaneously supports dynamic operations like deletion or insertion of files.

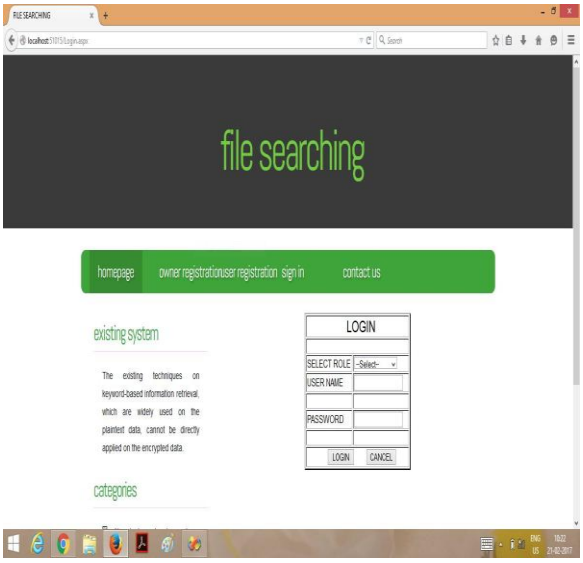


Fig 1. Login Page

Fig 2. File Upload Page

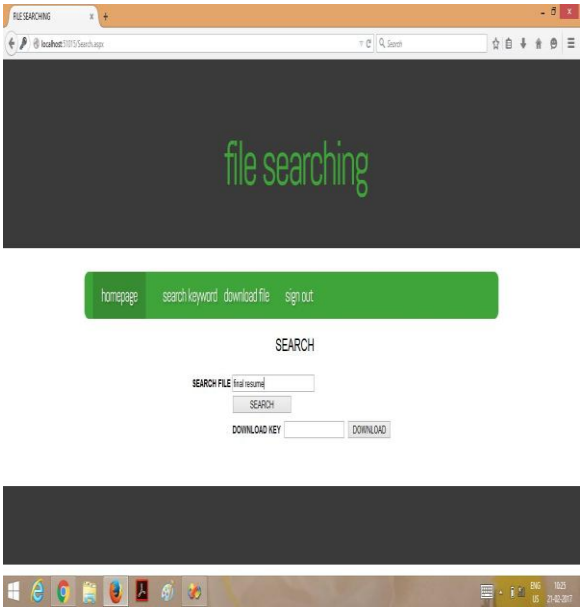


Fig 3. File Searching Page

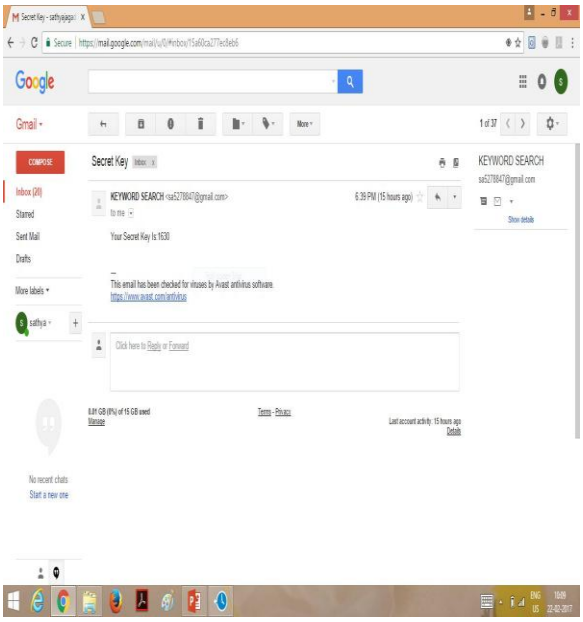


Fig 4. Keyword Generation to Registered Mail

VI CONCLUSION

In this paper, we formalize and solve the problem for the data users to download the encrypted files from the cloud server that has been hosted by the data owner. We designed two specific techniques that is SHA algorithm for the encryption of files and Fuzzy keyword algorithm for the searching process.

This system satisfies the needed requirements and this is the scope for the development of the project. The system is user friendly. Almost all the system objectives have been met and it satisifies the user needs. The system eliminates all the humen errors to zero level and it solves the problem that is being arised in the existing manual system. It has been tested completely and thus it overcomes the existing problems. The database design is flexible and it is more efficient. It satisfies all implemention and gone through all validation.

All phases of development are checked using different methodologies. The software is executed successfully and thus satisfies the objectives of the project. Further changes can be made in this system with minor modifications.

VII REFERENCES

[1] "Security challenges for the public cloud," published in 2012 by K. Ren, C.Wang, Q.Wang *et al.*
[2] "Cryptographic cloud storage," in *Data Security and Financial Cryptography* by S. Kamara and K. Lauter, in 2010.
[3]"A fully homomorphic encryption scheme," by C. Gentry,Stanford University, in 2009.
[4] "Software protection and simulation on oblivious rams," published in 1996 *by* R. Ostrovsky and O. Goldreich.
[5]"Public or asymmetric key encryption with keyword search," in *Advances in Cryptology-Eurocrypt by* D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano.
[6] "Practical techniques for searches on encrypted data," in *Security and Privacy, by* D. X. Song, D. Wagner, and A. Perrig
[7] "Secure indexes." *IACR Cryptology ePrint Archive*, by E.-J. Goh *et al.*, in 2003.
[8] "Searchable symmetric encryption" by R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky.
[9] "Fuzzy keyword search over encrypted data in cloud computing," was published in *2010 by* J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou.