International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 3 | ISSN : 2456-3307

# Fuzzy Keyword Search over Encrypted Data in Cloud

**Prof. Sonali Kale, Yash Gulati, Pooja Ghorpade, Avinash Savake**

Department of Computer Engineering, Savitribai phule University/Trinity Academy of Engineering, Pune, Maharashtra, India

## ABSTRACT

For privacy concerns, a secure search over encrypted cloud data has motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.

Hence, some cloud information retrieval schemes has been planned to resolve this drawback, these schemes not only will search the information they have properly, however can also stop sensitive information leaked out. During this paper, we tend to survey the privacy protective cloud information retrieval scheme and provide a comparison of them with regard to the key principles of search and privacy security.

**Keywords :** Fuzzy Keyword, Encrypted Data, RMSM, cloud servers, CSP

## I. INTRODUCTION

Cloud computing is a revolutionary technology that is changing the way IT hardware and software are designed and purchased. As a new model of computing, cloud computing provides abundant benefits including easy access, decreased costs, quick deployment and flexible resource management, etc. Enterprises of all sizes can leverage the cloud to increase innovation and collaboration. Despite the abundant benefits of cloud computing, for privacy concerns, individuals and enterprise users are reluctant to outsource their sensitive data, including emails, personal health records and government confidential files, to the cloud. This is because once sensitive data are outsourced to a remote cloud; the corresponding data owners lose direct control of these data. Cloud service providers (CSPs) would promise to ensure owners' data security using mechanisms like virtualization and firewalls. However, these mechanisms do not protect owners' data privacy from the CSP itself, since the CSP possesses full control of cloud hardware, software, and owners' data. Encryption on sensitive data before outsourcing can preserve data privacy against CSP. However, data encryption makes the traditional data utilization service based on plaintext keyword search a very challenging problem. A trivial solution to this problem is to download all the encrypted data and decrypt them locally. However, this method is obviously impractical because it will cause a huge amount of communication overhead. Therefore, developing a secure search service over encrypted cloud data is of paramount importance. Secure search over encrypted data has recently attracted the interest of many researchers. Song et al. first define and solve the problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that The main contributions of this paper are listed as follows:

- We will consider the problem of secure fuzzy keyword search.
- To generate dynamic key using fuzzy logic.

The CSP (Cloud Service Provider) may be a separate entity, thus cloud computing has several security problems, specially, the info security is extremely necessary and is also the foremost doubtless to threaten the customer's privacy. To protect the information, the data may be encrypted and out sourced to the cloud; however this may lead to a problem of information retrieval.

## II. METHODS AND MATERIAL:

**[1] J. ye [2016]:** This paper conveys some hassle for data search. Searchable secret writing permits users to look over the encrypted data on cloud storage to retrieve the concerned data while not coding. Throughout this paper, a fine-grained searchable scheme with a pair of non aforethought cloud servers is planned. During this paper, we have a tendency to propose a fine-grained searchable scheme supporting multiple users utilizing the advantage of attribute-based coding techniques

**[2] G. Arthi [2016]:** Multi keyword search mechanism is explains that the users can search among the cloud merely per their search. In proposed system new public-key cryptosystems is planned to be secure, efficiently, and easily share knowledge with others in cloud storage. The most set up is that one can mixture any set of secret keys and build them as compact collectively key, but all keys ought to be collective. This system is additional versatile than hierarchic key assignment. AES technique is employed within the projected system for effective data sharing.

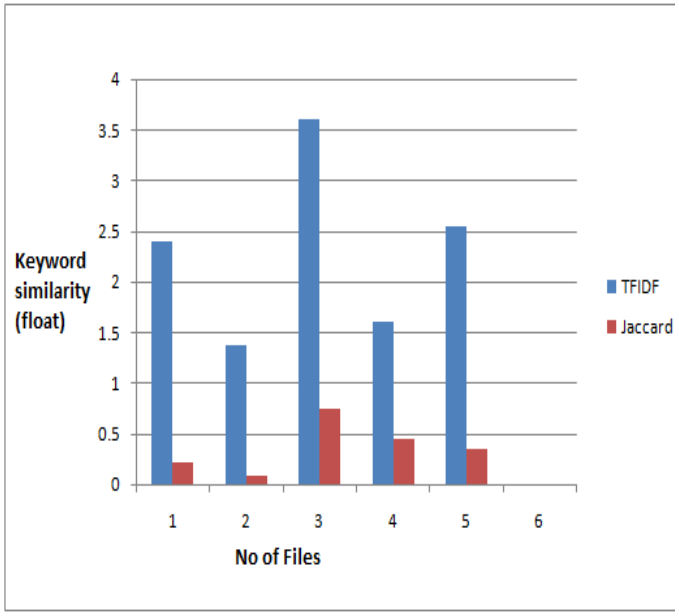
**[3] J. Shen [2015]:** This paper explains enhancement of the network and huge expansion of data. The data owner used to remotely outsources their data to cloud, which could avoid the native info management and scale back the native hardware worth. The supply encrypted info to cloud can increase the matter of the information retrieval, as a result of knowledge owner or unauthorized users can't search properly the data they need, and in addition it's impractical to transfer all of the information to native facet from the cloud that is in a position to guide to giant communication a computation overhead.

**4] W. Zhang [2014]:** In this paper, we have a tendency to propose schemes to touch upon secure hierarchic multi-keyword search during a multi-owner model. To modify cloud servers to perform secure search while not knowing the particular knowledge of each keywords and trapdoors, we have a tendency to consistently construct a unique secure search protocol. To rank the search results and preserve the privacy of relevancy scores between keywords and files, we have a tendency to propose a unique Additive Order and Privacy conserving perform family. To modify the cloud server to perform secure search among multiple owners' knowledge encrypted with totally different secret keys, we have a tendency to consistently construct a unique secure search protocol. Another resolution is to share a secret key among all knowledge house owners. However, this live can cause the protection threat of single purpose of failure. i.e., once the key secret's unconcealed by an information owner (e.g., careless key management), alternative knowledge owners' secret key are going to be leaked similarly. On the opposite hand, none people would be willing to share our personal secret keys with others in apply. To rank the search results and preserve the privacy of relevancy scores between keywords and files, we have a tendency to propose a unique Additive Order and Privacy conserving perform family.

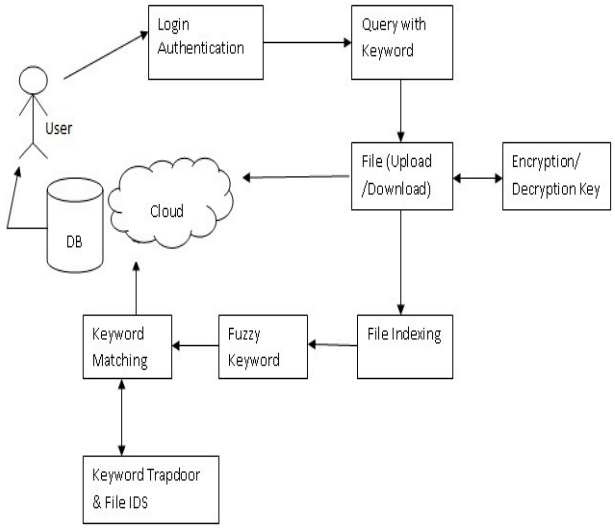
## III. RESULT:

The proposed work must provide full security to the search algorithm and searching over encrypted cloud data. Encryption rate of the searched keyword must more. The searching operation should be deploying to the cloud server. Searched keywords and all its related data must be ranked in specific format. The cloud server must calculate the total of encoded relevance scores and ranks them on the basis of this total. Fig. shows the encoding efficiency of our proposed AOPPF. We observe that PRMSM spends a little more time than SRMSM on trapdoor generation; the reason is that PRMSM introduces an additional variable to ensure the randomness of trapdoors. Fig shows the how to increases no of files size with respect to keyword similarity and TFIDF trapdoor vary rapidly, jaccard keyword similarity shows decreases slowly. As we can see from Fig. the more keywords existing in the cloud server, the more time is required for pairing operation. In this result TFIDF is better than JACCARD similarity.

So this type of result is shown in below graph of similarity of files.




## SYSTEM ARCHITECTURE




### 1. Authentication-secret key generation

In the authentication phase, user login to the system by providing his credentials. Then system authenticate user by verifying this credentials. Secrete key generated to give authenticated user. This secrete key generated by using hash function and secrete key generation algorithm.

### 2. Indexing

This is second module of the proposed system. Indexing is done on the uploaded and downloaded file. Indexing is done for file reference. For that, in the proposed system we have used context based indexing.

### 3. Encryption

Following are the few conditions which would be satisfied for encrypting keyword, first is data owners needs to utilized their own secrete key for encryption. Secondly, the secrete key must be encrypted to different cipher text every time for same keyword. These conditions are very advantageous to our proposed system for few reasons. First is trailing the secrete key of one owner wouldn't allow to disclose data of another owner. Second is cloud server would not look at any relationship between keywords which are encrypted.

### 4.Trapdoor calculation

The proposed system must satisfy following two conditions to make user of data to generate encrypted keywords (trapdoors) conveniently, efficiently and securely:

- First, the data user doesn't require asking to several data owners for secret keys to produce trapdoors.
- Second, each time the generated trapdoor must be different for same keyword. To meet these conditions, the generation of trapdoor is performed in two steps:
  Firstly, user of data produces trapdoor which is based on users search keyword as well as random number. Secondly, the trapdoors are re-encrypted by administrative server for authenticated users of data.

### 5.Top-k file display

The proposed system must fulfill conditions given next for ranking the relevance score while maintaining its privacy.

1) This function must save data order, as this helps cloud server for determining which file is more appropriate to a certain keyword, according to the encoded relevance scores.

2) Top-K function must not be exposed by the cloud server due to which cloud server can make comparison evaluation on encoded relevance scores without knowing their actual values.

3) Special data owners must have special functions such that illuminating the encoded data owner value

wouldn't results the leakage of encoded values of other data owners.

## IV. ALGORITHMS

1) AES Algorithm for File Encryption
2) TFIDF
3) Jaccard similarity Algorithm
4) Trapdoor Generation
5) Globally unique identifier (GUID) algorithm for file indexing.

## V. CONCLUSION

To enable the cloud server to perform secure search among multiple owners' data encrypted with different secret keys. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. This approach is efficient, even for large data and keyword sets.

## VI. REFERENCES


[1]. J. Ye, J. Wang, J. Zhao, J. Shen, K-C Li, "Fine-grained searchable encryption in multi-user setting," Soft Compute DOI 10.1007/s00500-016-2179-x, © Springer-Verlag Berlin Heidelberg 2016.

[2]. G. Arthi et.al, "Efficient search of Data in Cloud Computing using Cumulative Key," IJSTE - International Journal of Science Technology & Engineering, Volume 2, Issue 09, March 2016.

[3]. J. Shen et.al, "Privacy Preserving Search Schemes over Encrypted Cloud Data: A Comparative Survey," 2015 First International Conference on Computational Intelligence Theory, Systems and Applications.

[4]. W. Zhang et.al, "Secure Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing," 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks.

[5]. W. Zhang, Y. Lin, "Catch You if You Misbehave: Ranked Keyword Search Results Verification in Cloud Computing," Member, IEEE, JOURNAL OF LATEX CLASS FILES, VOL. 6, NO. 1, JANUARY 2015.