Privacy in Cloud Computing by Fuzzy multi-keyword search for Multiple data owners

Miss. Archana P. Tupkar[1], Prof. Varsha Dange[2]

[1]ME Student, Dept. of Computer engineering , DPCOE, Wagholi, Pune, Maharashtra, India
[2]Assistant Professor, Dept. Computer Engineering, DPCOE , Maharashtra, India

---***---

Abstract - *Cloud Computing is the technique which is used widely over the world. Data storing and retrieving are the main purpose to use cloud. So, it is necessary that cloud should be secured in any way. For privacy , there are various researches works under the single owner model but in fact cloud servers not only support single owner but also support multi owners to use the advantages provided by cloud computing. Sensitive data are necessary to be encrypted before it store in the cloud, usually cloud servers need to support feature of searching of keyword for these encrypted files. Fuzzy keyword search is used here to get approximate results. Until fuzzy keyword search is used only for single keyword but in this paper we are using fuzzy multi keyword search for more relevant results.*

Key Words: cloud computing, encryption, fuzzy keyword, multi owner, multi keyword.

1.INTRODUCTION

Due to the various benefits of cloud computing, various individuals and enterprises are interested in storing more sensitive data such as customers information, personal health records, emails, secret files of government to the cloud. Once sensitive data are employ to the remote cloud, related data owners lose direct control of data. By using virtualization and firewalls Cloud Service Providers(CSP) give owners data security but these techniques does not protect owners data privacy from CSP itself. To maintain data privacy against CSP, data should be encrypted before it stored on cloud. Data encryption process enable to make the data utilization service which is based on plaintext keyword search is a challenging problem and to solve this problem all encrypted data have to download and then decrypt them locally. This method is very impractical because it causes huge amount of communication problem. To overcome this problem it is important to develop a secure search service over encrypted data. This researches minimizes the computation and storage cost as well as secure keyword search over encrypted cloud data. There are various techniques for keyword search such as secure ranked multi keyword search, fuzzy keyword search. Sometimes users type or use slightly different formats e.g. cloud computing versus cloud-computing, Thus fuzzy keyword search is used to get approximate results. Mostly the techniques are used for single keyword but in this paper we are using fuzzy search for multi keywords for more relevant results. In this paper we construct a novel secure search protocol which is used to enable cloud servers to perform secure search without knowing the real value of both keywords and trapdoors. As a result, different keys are used by different data owners to encrypt the files and keywords. Authenticated data users can issue a query without knowing the secret keys of these different data owners.

1.1 Literature Survey

Due to advantages of cloud computing, it is become popular for data owners to store their confidential data on the cloud. For privacy, several researches which works under single owner model are motivated by secure searches over encrypted cloud data but in practical cloud servers not just support single owner model but also support multi owner model so Wei Zhang , Yapping Lin and some more proposes scheme to deal with Privacy Preserving Ranked Multi-keyword search in a Multi owner
model(PRMSM)[1].

To improve the accuracy of search result and to enhance user searching experience it is crucial for ranking system to support multiple keyword search, as a single keyword search often yields far too coarse results[2]. Co-ordinate matching i. e. as many matches as possible, is an efficient technique of multi keyword semantics to refine the result and has been widely used information retrieval (IR) system[3]. Traditional searchable encryption schemes typically only support exact keyword matches to overcome this problem, fuzzy keywords search is used to get approximate results .Recently some researchers propose wildcard based approach to provide fuzzy key word search[4].M Chuah , W. Hu exploit edit distance algorithm to quantify keywords similarity and designed two advanced techniques wildcard based and gram based techniques to construct storage efficient fuzzy keyword sets[6]. Their solution only provide single fuzzy keyword search e.g. searching for all data files that contain keyword having edit distance of 1 from the word architecture[5].

A user which is looking for files containing the word General system architecture will have to conduct three searches with the keywords General, system, architecture. Then user will have to decrypt the meta description of all the three returned list that contains all that three words General system architecture. In this work author proposes a privacy aware bedtree based solution that supports fuzzy multi-keyword search[7]. In the cloud computing users should be able to just use the cloud as if it is local, without worrying about the need to verify its integrity. Thus enabling public auditability for cloud storage is important so that users can resort to Third party auditor(TPA) to check the integrity of data and be worry free. Cong wang, kui Ren propose a secure cloud storage system which supports privacy- preserving public auditing. TPA is beneficial for cloud service providers to improve their cloud based platforms[9].

1.2 Problem Statement

To make easy to cloud servers to act safe search with zero knowledge of real data of both keywords and trapdoors to obstruct the attackers from to listen secretly to a private secret keys and to believe to be authenticated data users bow searches. We developing the Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM).To get approximate results we are using Fuzzy multi-keyword search with Edit distance algorithm.

2. IMPLEMENTATION DETAILS

2.1 Proposed System Architecture

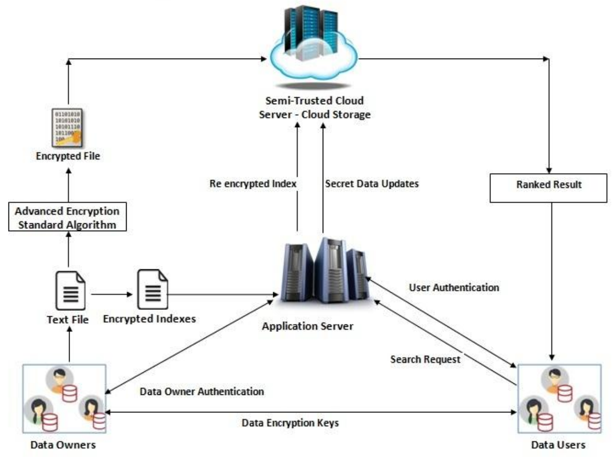

Fig -1: Proposed System Architecure

In this multi-user and multi-owner cloud computing model there are main four entities as shown in above fig 1. They are,

2.1.1Data owner:
Data owner consist of set of files which will be encrypted and to enable secure search data owners create index file and send that index file to the application server and data owner also send that encrypted files to the cloud server.
2.1.2 Application Server:
When index files are received, application server re-encrypt that index file for the authenticated data owners and send that re-encrypted file to the cloud server.

2.1.3 Cloud Server:
When trapdoors are received, the cloud server search the encrypted index of each data owner and returns the related set of encrypted files. Once data user receive the files, these files are decrypted by data user.

2.1.4 Data User:
Encrypted files are stored on cloud server and
when data user wants to search keywords by using normal search or fuzzy search over that files, data user computes related trapdoors and send that trapdoors to the application server. Once application server submits authentication to data user, then application server re-encrypt trapdoors and send it to the cloud server.

2.2 Mathematical model

In the above mathematical model there are four states as follows:-

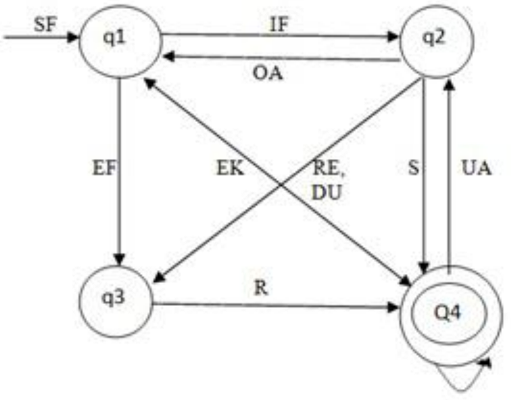

Fig -2: Mathematical Model

F=(q1,q2,q3,q4)
q1-Initial State
q4-Final State

and in this model there are various inputs and outputs
Where,
SF = set of files
IF = indexing file
OA = data owner authentication
EF = encrypted files
EK = encryption keys

RE = re-encrypted index
DU = secret data updates
UA = user authentication
S = search request
R = Ranked result
q1 -Data owner
q2-Application Server
q3- Cloud Server
q4-Data User

3. ALGORITHMS

3.1 Edit Distance Algorithm:

We can calculate edit distance between two string with the help of their end we have to work on their ends .The three steps are as follows:-

Step 1. At the end of string add one character.
Step 2. Delete one character from the end of string.
Step 3. Substitute the character at the end of string.

For ex- Edit distance between TREE and FREE is 1. Edit distance between NETWORK and NETWORKING is 3.

3.2 AES Algorithm

The schematic of AES structure is given in the following illustration:- AES consist of many rounds and each round consist of four steps as follows:-

Step 1. Byte substitution(Sub Bytes)
Step 2. Shift rows
Step 3. Mix columns
step 4. Add round key

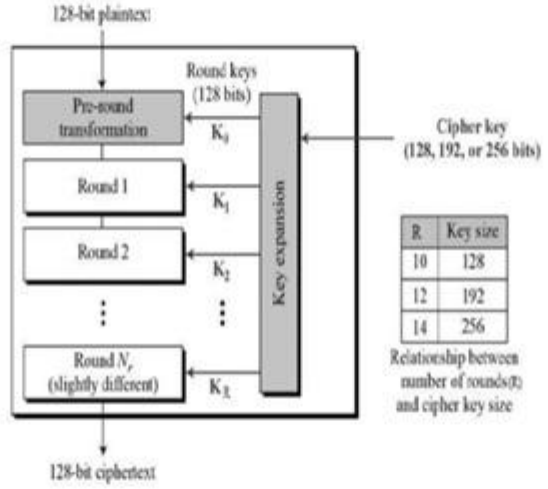

Fig -3: AES Structure

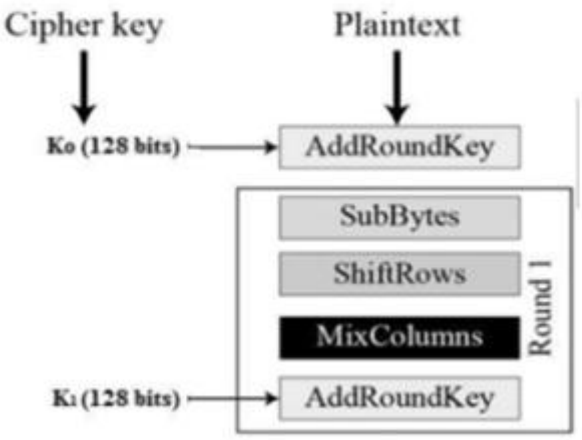

Fig -4: Operations in AES in 1st round

4. EXPERIMENTAL RESULTS

4.1 Normal Search vs Fuzzy Search

X axis- keywords

Y axis- No. of files we get

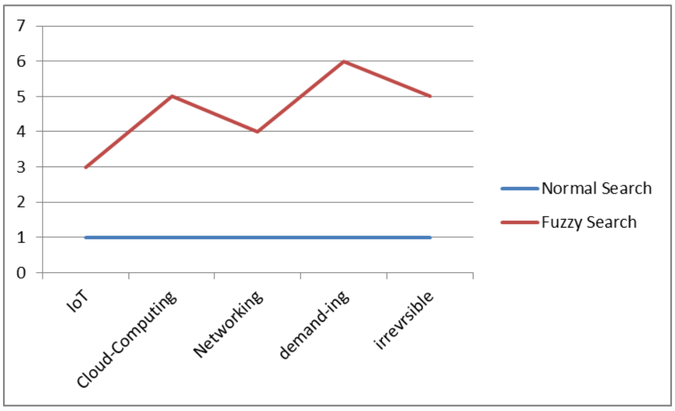

Chart -1: Normal Search vs Fuzzy Search

3. CONCLUSIONS

In this paper we solve the problem of multi keyword search by multiple owners. To achieve secure, efficient searches over multiple data owners our proposed scheme uses authenticated users. We used fuzzy keyword search for more relevant results.

ACKNOWLEDGEMENT

Special thanks to authors Wei Zhang, Sheng Xiao, Yaping Lin, Jie Wu, and Siwang Zhou for valuable existing work in this area. I am thankful to my Guide Prof. Dange Varsha and Head of Department Prof. Dandavate Arati, Contributed to this paper for their valuable comments and sharing their knowledge and idea.

REFERENCES

[1] Wei Zhang, Yaping Lin, Sheng Xiao, Jie Wu, Siwang Zhou,Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing ,IJCST ,Volume 2 Issue 3, pp.60.-64,June-2015.

[2] C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, Privacy preserving public auditing for secure cloud storage, International Journal of Advances in Science and Technology (IJAST),Vol 2, Issue 4 ,pp.41-44,December 2014.

[3] M. Chuah, W. Hu, Privacy-aware BedTree Based Solution for Fuzzy Multi-keyword Search over Encrypted Data , SSRG International Journal of Mobile Computing Application, volume 2, Issue 3 ,pp.38-44,June 2013 .

[4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, A view of cloud computing, Communication of the ACM , vol. 53, no. 4, pp. 5058, 2010.

[5] D.Song, D.Wagner, and A.Perrig, Practical techniques for searches on encrypted data , in Proc. IEEE International Symposium on Security and Privacy (SP00), Nagoya, Japan,pp. 4455 Jan. 2000.

[6] D. B. et al., Public key encryption with keyword search secure against keyword guessing attacks without random oracle , EUROCRYPT, vol. 43, pp. 506522, 2004.

[7] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, Privacy preserving multi-keyword ranked search over encrypted cloud data, Parallel and Distributed Systems , IEEE Transactions on,vol. 25, no. 1, pp. 222233, 2014.

[8] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, Verifiable privacy preserving multi-keyword text search in the cloud supporting similarity-based ranking, Parallel and Distributed Systems, IEEE Transactions on,vol. 25, no. 11, pp. 30253035,2014.

[9] S. Bhati,A. Bhati,S.K Sharma, A New Approach towards Encryption schemes: Byte- Rotation Encryption Algorithm. World CECS,Vol-2, pp.24-26, 2012.

[10] P. Xu, H. Jin, Q. Wu, and W. Wang, Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack, Computers IEEE Transactions on,vol. 62, no. 11, pp. 22662277, 2013.

[11] S. S. Khan, R.R. Tuteja, Security in Cloud Computing using Cryptographic Algorithms, IJCA,Vol. 3, Issue 1, ISSN 2320- 9798,pp 148-157 January 2015.

[12] Shelna Valsan K.P,Varshap, Enhancing Cloud Security and Integrity by Using multiple Encryption Algorithms and Stripping, IJSR, IJCA,Volume 4 Issue 4, ISSN 2319-7064,pp 1065- 1068 ,April 2015.