ENCRYPTED CLOUD DATA WITH A SECURE AND DYNAMIC MULTI-KEYWORD RANKED SEARCH

MANJUNATHA K J[1], PRAKASH[2], DEVKI P [3]

[1]UG STUDENT, DEPT.OF IS&E, THE NATIONAL INSTITUTE OF ENGINEERING, MYSURU, INDIA

[2]UG STUDENT, DEPT.OF IS&E, THE NATIONAL INTITUTE OF ENGINEERING, MYSURU, INDIA

[3]ASSOCIATE PROFESSOR, Dept. of IS&E, THE NATIONAL INSTITUTE OF ENGINEERING, MYSURU, INDIA

Abstract - *Due to the increasing popularity of cloud computing, many data owners upload their data to cloud servers for great convenience and reduced cost in data management. However, responsive data should be encrypted before uploading for privacy requirements, which obsoletes data utilization like keyword-based document search. We present a secure multi-keyword ranked search scheme over encrypted cloud data, which supports dynamic update operations like deletion and insertion of documents.*

***Key Words*:Encryption; Multikeyword; Ranked search**;

1. INTRODUCTION

In the cloud storage, protecting data privacy is a big challenge for data owners.A traditional way to solve this problems is to encrypt documents before uploading documents to cloud server.However, this method is dramatically degardes the performance of text search on encrypted documents.We propose an architecture of multikeyword ranked search over encrypted documents. Their techniques provide provable secrecy for encryption , in the sense that the untrusted server cannot learn any thing about the plaintext given only the cipher text.

Both data owner and users are impressed to uploading their data to the cloud server, instead of purchasing software and hardware to manage the data. Multi-keyword ranked search achieves more and more attention for its practical applicability. Recently, some dynamic schemes have been proposed to support inserting and deleting operations on document collection

2. LITERATURE REVIEW

Single keyword search over encrypted data on cloud:

Taking this searchable encryption scheme provide to user to confidently look for over encrypted data by keyword without first applying decryption on it,the proposed techniques support only Boolean keyword search,without capturing any applicability of the files in the search result.

Multi-keyword search over encrypted data on cloud:

In this search for known cipher text model and background model over encrypted data providing low computation and communication overhead. The coordinate matching is chosen for multikeyword search.We used inner similarity for ranking files. Ranked search returning the matching files in a ranked order regarding to certain important criteria.

2.1 PROPOSED SYSTEM

We propose a secure tree-based search scheme over the encrypted cloud data, which supports multi-keyword ranked search and dynamic operation on the document collection. Specifically, the vector space model and the widely-used "term frequency (TF) × inverse document frequency (IDF)" model are combined in the index construction and query generation to provide multi-keyword ranked search.

In this, the system refers to three entities,as illustrated in fig 1,the data owner, the data user, and the cloud server.the data owner is responsible for collecting documents F,building hierarchical clustering index and uploading them in an encrypted format to the cloud server. Data owner also needs to get the authorization to the data users. The cloud server provides a huge storage space, and the computation resources needed by ciphertext search.

3.PROBLEM FORMULATION

Notation:

- Nf;wi – The number of keyword wi in document f•
- N – The total number of documents •
- Nw – The number of documents that contain key-word wi•
- TFf;w I – The TF value of wi in document f •
- IDF′wi – The IDF value of wi in document collection.
- TFu;w – The normalized TF value of keyword wi stored in index vector D u.•
- IDF wi – The normalized IDF value of keyword wi in document collection.

The relevance evaluation function is defined as:

$$RScore(Du ; Q) = Du \cdot Q = \sum wi \in Wq\ TF\ u ; wi \times IDF\ wi : \quad (1)$$

If u is an internal node of the tree,TF u;wi is calculated from index vectors in the child nodes of u . If the u is a leaf node, TF u;w I is calculated as:

$$TF\ u;wi = TF'f;w / i\sqrt{\sum wi \in W(TF'f;wi)2}; \quad (2)$$

Where TF′f;wi = 1 + ln N f;wi . And in the search vector Q,IDF wi is calculated as:

$$IDFwi = IDF'wi / \sqrt{\sum wi \in Wq(IDF'wi)2}; \quad (3)$$

Where IDF′ wi = ln(1 + N / N wi).

Keyword Balanced Binary Tree.The balanced binary tree is widely used to deal with optimization problems. The keyword balanced binary (KBB) tree in our scheme is a dynamic data structure whose node stores a vector D . The elements of vector D are the normalized TF values. Sometimes, we refer the vector D in the node U to D u for simplicity. Formally, the node u in our KBB tree is defined as follows:

$$U = \langle ID ; D; P\ l ; Pr ; FID \rangle; \quad (4)$$

Where ID denotes the identity of node u , P l and P r are respectively the pointers to the left and right child of node u .

consisting of the TF values which is calculated as follows:

$$D [I] = max \{ u:P\ l \rightarrow D [I] ; u:Pr \rightarrow D [I] \} ; I = 1 ; :::; m: \quad (5)$$

4. SYSTEM DESIGN

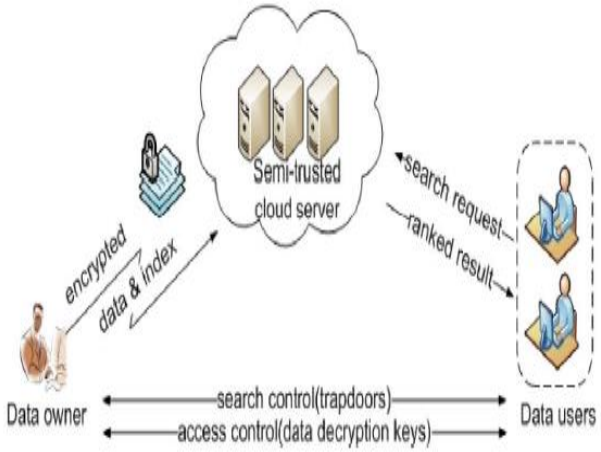

Fig.1

Data Owner Module:

This module helps the owner to register those details and also include login details. This module helps the owner to upload his file with encryption using RSA algorithm. This ensures the files to be protected from unauthorized user.

Data User Module:

This module includes the user registration login details. This module is used to help the client to search the file using the multiple key words concept and get the accurate result list based on the user query

Cloud Server and Encryption Module:

This module is used to help the server to encrypt the document using RSA Algorithm and to convert the encrypted document to the Zip file with activation code and then activation code send to the user for download.

Rank Search Module:

These modules ensure the user to search the files that are searched frequently using rank search. This module allows the user to download the file using his secret key to decrypt the downloaded data. This module allows the Owner to view the uploaded files and downloaded.

5. CONCLUSIONS

In this survey we studied different cryptographic techniques for data sharing security. One trivial solution is proposed, a secure, efficient and dynamic search scheme, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents. We construct a special keyword balanced binary tree as the index,and propose a "Greedy Depth-first Search" algorithm to obtain better efficiency than linear search. In addition, the parallel search process can be carried out to further reduce the time cost.

REFERENCES

[1] K. Ren, C.Wang, Q.Wang et al., "Security challenges for the public cloud," IEEE Internet Computing.

[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010.

[3] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.

[4] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir

queries," in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50–67.

[5] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on . IEEE, 2000, pp. 44–55.

[6] E.-J. Gohet al. , "Secure indexes." IACR Cryptology ePrint Archive,vol. 2003, p. 216, 2003.