International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue VII, July 2017- Available at www.ijraset.com

Multi-Keyword Top-K Ranked Search over Encrypted Cloud Using Parallel Processor

Pawar Supriya[1], Dr. S. A. Ubale[2]
[1,2]*Department of Computer Engineering, Zeal College of Engineering and Research, Pune. Savitribai Phule Pune University*

Abstract: ***In this paper, we introduce a secure multi keyword Ranked Search over encrypted cloud information which at the same time supports dynamic upgrade operation like deletion and insertion of reports detail the vectored space model and the widely utilized TF-IDF (Term Frequency - Inverse Document Frequency) model are combined in the index construction and query generation. Due to the increasing popularity of the cloud, more and more data owner are motivated to outsource their data to cloud server for great conveniences and reduced cost in data management. The data should be in encrypted form before outsourcing for privacy requirement. In particular, the vector space demonstrates and the broadly utilized TF IDF models are combined in the index construction and query generation. We develop a special tree-based index structure and propose a Greedy Depth-first Search algorithm to give productive multi-keyword Ranked Search..***
Keywords: ***Cloud Computing, Searchable Encryption, Multi-Keyword Ranked Search, Dynamic Update.***

I. INTRODUCTION

Cloud computing has been considered as another model of enterprise IT foundation, which can organize huge resource of processing, storage and applications, and empower clients to enjoy ubiquitous, helpful and on request network access to a shared pool of configurable computing resources with extraordinary effectiveness and minimal economic overhead Attracted by these appealing features, both people and enterprises are motivated to outsource their information to the cloud, instead of purchasing software and hardware to [9]deal with the information themselves. Despite of the different ad-vantages of cloud administrations, outsourcing sensitive data, (such as e-mails, personal health records, company finance data, government documents, etc.) to remote servers brings security concerns. The cloud specialist providers (CSPs) that keep the information for clients may get to clients' sensitive data without approval. A general way to deal with secure the information secrecy is to encode the information before outsourcing. However, this will bring about a huge cost as far as information usability. For Example, the current procedures on keyword based data retrieval, which are generally utilized on the plaintext information, can't be directly applied on the encrypted information. Downloading the all types of data from the cloud and also decrypt locally is obviously impractical [5, 6, and 7].

In order to address the above problem, researchers have designed some general-purpose solutions with fully homomorphice encryption. However, these methods are not practical due to their high computational overhead for both the cloud sever and user. On the contrary, more practical special purpose solutions, such as searchable encryption (SE) schemes have made specific contributions in terms of efficiency, functionality and security.

Searchable encryption schemes enable the client to store the encrypted data to the cloud and execute keyword search over cipher text domain [9,10]. So far, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity search, multi-keyword Boolean search, ranked search, multi-keyword ranked search, etc. Among them, 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. These are significant works as it is highly possible that the data owners need to update their data on the cloud server. But few of the dynamic scheme support efficient multi keyword ranked search.

II. LITERATURE SURVEY

K. Ren, C. Wang,et. al. [1], focus on the first encryption scheme based on search-able paradigm and the search time is proportional to the size of the data. The size of the data if changed or increased led to the change in the time required for running the encryption algorithm. Other proposed wor3k includes two schemes for encryption data. The proposed scheme has single keyword search mechanism which has adaptive and chosen keyword attacks. This single keyword search mechanism shows limitations

W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, et. al. [2], propose Multi-keyword search mechanism users have to put multiple keywords. Based on the keywords given as an input to the system, the documents are retrieved. The various flavours of multi-keyword includes conjunctive keyword search where all the documents are outputted by the system. The documents which are outputted by the system have all the keywords given as an input by the users. The next keyword mechanism includes disjunctive approach of document retrieval where the documents are retrieved based on the subset off attributes submitted by the users. The documents which are outputted have few or all the attributes submitted by the users. The combination of conjunctive and disjunctive approach shows the combined result of all the keywords inputted by the user to the system. The only limitation in this multi-keyword approach is that the ranking of the documents are missing. [8] Few of the documents which are irrelevant are topped in the showcase whereas few of the documents which are relevant are not topped in the retrieval result.

N. Cao, C. Wang, M.et. al. [3],uses Ranked search algorithm which is shows Ranked search can be used to implement the retrieval of data from the cloud with the help of relevancy. In this approach, the top relevant k attributes are searched, fetched and listed in the output. The main disadvantage of this system is that the ranked algorithm can only work fine with single keyword search and cannot work well with multiple keyword search strategy. The new approach that is proposed further is privacy preserving which considers the matching phenomenon in matching the most relevant keywords with the documents. However the drawback of this system is that the matching is done with linearity and it is not done considering the importance of a particular keyword over the other keyword.

W. Zhang, S. Xiao,et. al. [4], new theory is proposed which has secure multi-keyword search engine which has similarity based ranking and thus includes TF, IDF evaluation. This algorithm is better than the previous algorithm however it shows precision loss. The new theory proposed further has indexed mechanism thus making the searching easier. In the previous approach, the whole document where searched for the relevant keyword matching. When the document is updated slightly or completely, entire document has to be searched to identify the relevant information. However this problem is solved with the help of indexing where the documents indexes has most relevant keywords associated with the document. The entire is not searched for when the user puts certain keywords to fetch the document. However the new problem that has arisen due to indexing is false positives.

Ubale S. A, Dr.Apte S S, et. al. [8] focus on security issues in cryptographic role-based access control system. In this paper it secures data storage in cloud using most popular algorithm that are RSA and AES. Using these algorithms any organization can easily upload the data in public cloud. It uses experience trust model which is used for securing organization data. Result of encryption and decryption keeps constant size of ciphe text and decryption key.

III. PROPOSED SYSTEM

This paper propose a secure tree based search scheme over the encrypted cloud data which supports multi keyword ranked Search and dynamics operation on the document collection. The Secure KNN Algorithms is utilized to encrypt the index and query Vectored and means while ensure accurate relevance score calculation between encrypted index and query vectors. We design a searchable encryption scheme that supports both the flexible dynamic operation on documents collection. Due to the tree based index the search complexity of proposed scheme is fundamentally kept to logarithmic and in practice the proposed scheme can achieve higher search efficiency by executing our “Greedy Depth First Search” Algorithm. We Define and Solve the challenging problem of “Privacy preserving multi keyword ranked Search over encrypted cloud data” and establish a set of Strict privacy requirements for such a secure cloud data utilization system to become reality. Among various Multi keyword Semantics we chose the effects principal of matching query.

Fig. 1. Architecture of Ranked Search

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue VII, July 2017- Available at www.ijraset.com

Ranked search can be used to implement the retrieval of data from the cloud with the help of relevancy. In this approach, the top relevant k attributes are searched, fetched and listed in the output. The main disadvantage of this system is that the ranked algorithm can only work fine with single keyword search and cannot work well with multiple keyword search strategy. The new approach that is proposed further is privacy preserving which considers the matching phenomenon in matching the most relevant keywords with the documents. However the drawback of this system is that the matching is done with linearity and it is not done considering the importance of a particular keyword over the other keyword.

A. Data Owner

Data owner is the person who has the data to be published on to the cloud. He is the person who wants to keep data on the cloud for further access. He first builds the tree structure of all the attributes. The attributes are the keywords which are relevant to the documents uploaded by him on to the cloud. He uploads the set of documents which has the data and the index tree structure which is used to access the data from the cloud. He also distributes the access policy to determine the access control. He also shares the key information to the data users to download the data once the data is accessible to the user based on the access policy.

B. 2. Data User

2. Data Users – Data Users are eager to read the documents placed on the cloud. Today most of the information is stored on the cloud. The data users access the data from the cloud using query keywords. According to the access control, they are checked for the query that they propose. Once the access control is verified, users are allowed to decrypt the data.

C. Cloud Server

Cloud server is the manager of the entire cloud man-agement. Cloud executes the tree structure and sends the most significant search records to the user. Once the user updates the documents, the cloud server takes up the responsibility of updating the cloud information and the tree structure.

IV. ALGORITHMS

A. Build Index Tree

*1) Input:*Collection of documents F = f1, f2, …….. fnusing identifiers FID = FID | FID = 1, 2, ……10

*2) Output:*we get index tree.

B. Index Tree

Node u if the node u is not a leaf node then

```
{
 if RScore (Du,Q) > k^thScore then
goto;
greedy depth first search of higher relevance child
greedy depth first search of lower relevance child
else
return
end if
else
if RScore (Du, Q) < k^thScore then
Delete element which have small score
And insert the new element
end if
return
end if
 }
```

C. Process

1) Input : Encrypted Files: System's input

2) Output: Ranked Result.

3) Process:

Step 1: Data owner Select the File.
Step 2: Encrypt File (For encryption we use AES and RSA Algorithm).
Step 3: Upload the file on Cloud Server.
Step 4: Data Owner generate secret key for file.
Step 5: Data Users Searches the file and send request to the data user.
Step 6: Data Owner responses to user and sends secret key on E-mail.
Step 7: Data User Receive the secret key.
Step 8: Finally Data User successfully gets downloaded file.
Step 9: After downloading ,Data user re-encrypted the file successfully.

V. PERFORMANCE MEASURE

Given graph x axis Precision and y axis is the number of result document is search result we get.

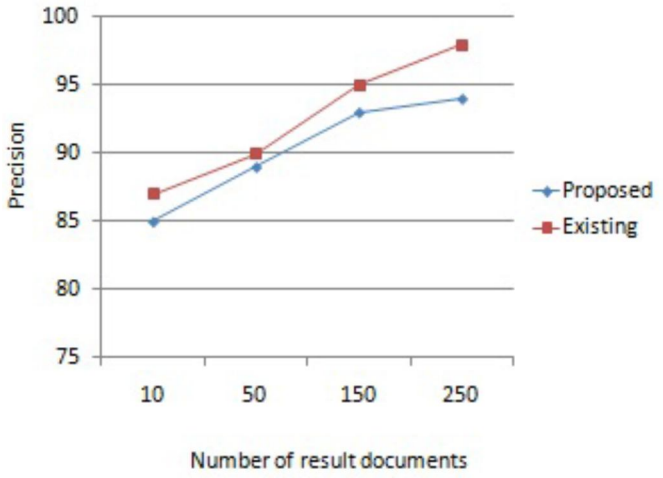

Fig. 2. Precision Analysis

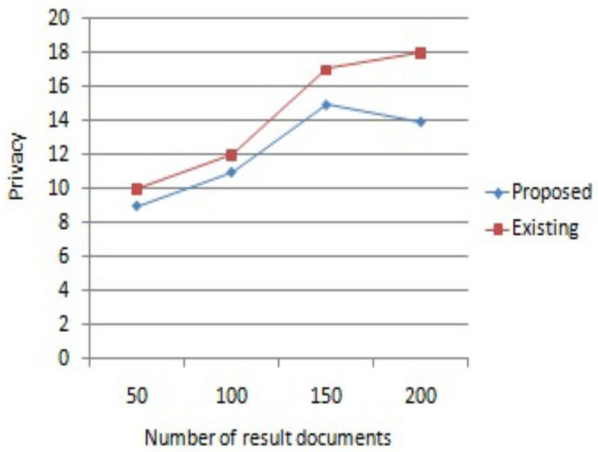

Fig. 3. Privacy Analysis

VI. SYSTEM ANALYSIS

Index tree construction process can be used for the document collection F. Two main steps are used in index tree Constructions are as follows:

A. Construct an unencrypted KBB tree based on document collection F.

B. Encrypting the index tree with separation and two multiplication of matrix called (m _m): The proposed search scheme can be executed in parallel, which further improves the search efficiency. For example, we assume there are a set 3of processors P= P1, . . . Pn available. Given a search request, an idle processor Pi is used to query the root r. If the search could be continued on both the children, and there is an idle 3processor pj, the processor pi continues to deal with one of the children while processor Pj deals with the other one. If there is no idle processor, the current processor is used to deal with the child with larger relevance score, and the other child is put into a waiting queue. Once there is an idle processor, it takes the oldest node in the queue to continue the search.

VII. CONCLUSION

This paper has proposed a dynamic approach to search the keywords and display the documents on the screen as the Retrieval result. The proposed work has multiple keyword search mechanism with ranked approach. The index tree is Constructed with the approach of balanced binary tree as the index. The algorithmic approach used in the paper has greedy approach which is most efficient as compared to the existing linear approach. Parallel search mechanism makes the execution of searching operation parallel thus ensuring simultaneously access. Secure KKN algorithm has been used in the paper to ensure security of the system. As the index is secure, the data owners are not kept active and they can go offline once they upload the data and submit the tree index structure of data access. It is very difficult to revoke the user as it is necessary to build the index once again. Hence it is necessary to forward the user management to the most active element in the cloud management which is the cloud server.

REFERENCES

[1] K. Ren, C. Wang, and Q. Wang, “Security challenges for the public cloud.“ IEEE Internet Comput., vol. 16, no. 1, pp. 6973, Jan-Feb. 2012.
[2] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li,“Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking,” in Proc. 8th ACM SIGSAC Symp. Inf., Comput. Commun. secur. , 2013, pp. 7182.Stanford Univ., Stanford, CA, USA, 2009.
[3] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” in Proc. IEEE INFOCOM , Apr. 2011, pp. 829837..
[4] W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, “Secure ranked multi-keyword search for multiple data owners in cloud computing,“ in Dependable Syst. Networks (DSN), IEEE 44th Annu.IEEE/IFIP Int. Conf. , 2014, pp. 276286.
[5] S. Kamara and K. Lauter, “Cryptographic cloud storage, in Proc. Financ Cryptography Data Secur., 2010, pp. 136149.
[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, “Public key encryption that allows pir queries,” in Proc. Adv. Cryptol., 2007, pp. 5067.
[7] S. Kamara and C. Papamanthou, “Parallel and dynamic search- able symmetric encryption,” in Proc. Financ. Cryptography Data Secur., 2013, pp. 258274.
[8] Ubale S. A, Dr.Apte S S, Bokefode J, Dr.Karande K.J, “Developing Secure Cloud Storage System by applying AES and RSA Algorithm with Role Based Access Control Model”, International Journal of Computer Applications, Volume 104 No.5, May 2015.
[9] Ubale Swapnaja A, Dr. Apte S. S, Bokefode Jayant D, , “Developing Secure cloud storage system by using Access Control Models“, ICDECT, Springer AISC Series March 2016.
[10] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, “Privacy preserving multi-keyword text search in the cloud sup- porting similarity based ranking, in Proc. 8th ACM SIGSAC Symp. Inf., Comput. Commun. secur. , 2013, pp. 7182.