Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 5, May 2017, pg.90 – 94

A Review Paper on Implementation of a Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data

A.P. Jaware, N.R. Borkar
Computer Science and Engineering .Amravati University INDIA
anita.jaware@gmail.com; namrata.borkar@gmail.com

Abstract*— The major aim of this paper is to solve the problem of multi-keyword ranked search over encrypted cloud data at the time of protecting exact method wise privacy in the cloud computing concept in this paper we are using three main parts first one is Data Owner, second one is Data User and third one is cloud computing. In this paper we introduce Implementation of a secure multi-Keyword Ranked search scheme over encrypted cloud data. Specifically, the vector space model and the broadly utilized TF×IDF model are joined as a part of the record development and query generation. We build a unique tree-based file structure and propose "greedy Depth- first Search" calculation to give efficient multi-keyword ranked search with no. of count matching multi-keywords. The protected KNN calculation is used to encrypt the file and query vector. The vector space model facilitates to offer enough search accuracy and the Triple DES encryption used to provide security.

***Keywords*— searchable encryption, multi-keyword ranked search, dynamic update, cloud computing**

I. Introduction

Cloud computing is one way of computing. Here the computing resources are shared by many users. The benefits of cloud can be extended from individual users to organizations. The data storage in cloud is one among them. The virtualization of hardware and software resources in cloud nullifies the financial investment for owning the data warehouse and its maintenance. Many cloud platforms like Google Drive, iCloud, SkyDrive, Amazon S3, Dropbox and Microsoft Azure provide storage services. Security and privacy concerns have been the major challenges in cloud computing. The hardware and software security mechanisms like firewalls etc. have been used by Cloud Service provider.

Data Owner keep their data to the cloud server and cloud server store encrypted data files that can be accessed by authorized data user. This Paper suggests a secure multi keyword ranked search scheme over encrypted cloud data which support Dynamic operation. Specifically the vector space model that widely used Term Frequency(TF)*inverse document frequency model are combined in index construction.

The effective data retrieval need, the large amount of documents demand the cloud server to perform result relevance ranking, instead of returning undifferentiated results with count it will display how many keywords matching with existing query search. Such ranked search system enables data users to find the most relevant information quickly, rather than burdensomely sorting through every match in the content collection,

Effectiveness, we develop a tree based list structure and based on this tree list we propose a “Greedy Depth –first search” calculation.

To oppose distinctive attacks in different threat models, we build two secure search schemes: the basic dynamic multi-keyword ranked search (BDMRS) Scheme in the known cipher text model, and the enhanced dynamic multi-keyword ranked search (EDMRS) Scheme in the known background model. Our commitments are condensed as takes after:
1) We plan a searchable encryption scheme that assist both multi-keyword ranked search and Dynamic operation on document collection.
2) The suggest scheme is support to logarithmic for search complexity in Structure of tree-based index. the proposed scheme can provide higher search functionality by executing our "Greedy Depth-first Search" algorithm. Additionally it also reduce the time cost of search process.

II. LITERATURE SURVEY

Cloud computing is an emerging computing technology that uses Internet and central Remote server to maintain documents and applications .It is Distributed computing on internet and delivering computing service on internet. Be that as it may, as the cloud's state Processing is rising and growing quickly both theoretically and actually, the legitimate/contractual, monetary, Administration quality, inter-operability, security and protection issues still posture critical difficulties. In this Part, we depict different services and organization models of distributed computing and recognize significant:

A. Searchable encryption schemes: Searchable encryption schemes (SES) able for the clients to store the encrypted data to the cloud and execute keyword search over cipher text domain. There are different types of cryptography primitives searchable encryption can be build using symmetric key based cryptography [7], [8], [9], [10].
1. The first symmetric searchable encryption (SSE) scheme, suggested by Song et al. [7] and the search time of their scheme is linear to the size of the data collection.
2. Regarding security issue for SSE, Goh [8] advice formal security definition and designed a scheme for SSE based on Bloom Filter. Searching time for this scheme is O(n), n is the cardinality of the document collection.
3. Curtmola et al. [10] advice two schemes that attain the optimal search time
  a. **SSE-1**: This scheme is secure against chosen-keyword attacks (CKA1)
  b**. SSE-2**: This scheme is secure against adaptive chosen-keyword attacks (CKA2)
These all works are based on single keyword Boolean search schemes that are very simple related to functionality.

B. **Threat Models** :When there are large scale of works is available then that works proposed under different thread module to achieve several search functionality such as **single keyword search**, similarity search[11], [12], [13], [14], **multi-keyword Boolean search** [15], [16], [17], [18], [19], [20], [21], [22], **ranked search** [23], [24], [25], and **multi-keyword ranked search** [26], [27], [28], [29], etc.
1. **Multi-keyword Boolean search** let the user to input multiple query keywords to request appropriate documents these will return documents that contain all of query keywords which cannot provided ranking functionality.
2. **Ranked search** can able for quick search of the most relevant document returning top-k most relevant document .but these are design only for single keyword search.
3**. Cao et al**[26] successfully completed the first privacy preserving **multi-keyword ranked search**.
the document are ranked according to number of matched query keywords. But it is not accurate enough and search efficiency is linear with the cardinality of the documents collection.
4. **sun et al** [27] show secure multi-keyword ranked search scheme. The authors build searchable index tree based on vector space module and measure by using term frequency and inverse document frequency (TF*IDF).Search efficiency is better-than-linear search but result in precision loss.
5**. Orencik et al**. [28] secure multi keyword search method which used Local sensitive hash LSH functions to bunch the similar documents but LSH cannot provide exact ranking.
6. **Zhang et al** [29].Proposed secure multi-keyword ranked search scheme in a multi-owner model. Proposed a scheme to deal with secure multi-keyword ranked search in a multi-owner model .This scheme don't support dynamic operations.
In this paper we can search multi-keyword ranked search with dynamic update operation like deletion.
Firstly in this paper required security so it provided encryption process for that it uses triple DES algorithm which provided security data can access by authenticated user.
7. **Role of data owner** is to check wether user is authenticated or valid if yes then owner gives permission to the authenticated user to upload and download particular document to the cloud and dynamic operation is also perform by data owner like deletion of documents.
8. Role of data user is to upload and download document to the cloud server and it can search document with multiple query as input and this search operation provided numbers of document with ranked as output top-k ranked provided most relevant document with count which provided how many keyword match in this documents ranked provided most relevant document with count which provided how many keyword match in this documents.

9.Role of cloud server: Cloud server stores the encrypted document collection C and the encrypted searchable tree index I for data owner .In the wake of tolerating the trapdoor TD from the data user, look over the index tree I, in conclusion gives back the relating gathering of top-k situated encoded reports. Also, in the wake of tolerating the update information from the data owner, the server needs to update the index I and document gathering C as per the received information

III.DESIGN GOALS

To find ranked search for effective utilization of outsourced cloud data under the aforementioned model, our System design should simultaneously achieve security and performance guarantees as follows.

1. **Multi-keyword Ranked Search:** To implements search schemes which access multi-keyword query and provide result similarity ranking for effective data retrieval, instead of returning number of document it will return exact documents with ranked.

2. **Privacy-Preserving:** To forbid the cloud server from acquisition additional information from the dataset and the index, and to provide privacy requirements specified

3. Efficiency: This also perform privacy should be achieved with low communication and computation overhead.

IV.PROPOSED SYSTEM

Cloud Data build index structure that is based on tree and for the purpose of delivery of appropriate multi-keyword ranked search with count that display how many keywords match with existing search query. The proposed system can realize sub-linear search time and deal with the deletion and insertion of documents flexibly. Wide -ranging experiments are shown to demonstrate the efficiency of the proposed scheme.

- Copious works have been suggest under different wild models to perform action on various search functionality,
- They are also perform some dynamic operation like deletion operation on Text Documents.

In this paper we are used to search multi keywords with ranked functionality. We can also fount count functionality which can tell us how many keywords will be matching with existing search query.

For searching Purpose we are using TF*IDF module.

V. SCOPE OF PROJECT

Given system built a special structure that used Triple DES algorithm for encryption and suggest a “GDFS” algorithm to offer effective multi-keyword search. The advice system can realize sub-linear search time and deal with the deletion documents openly. Extensive experiments are conducted to demonstrate the efficiency of the proposed system.

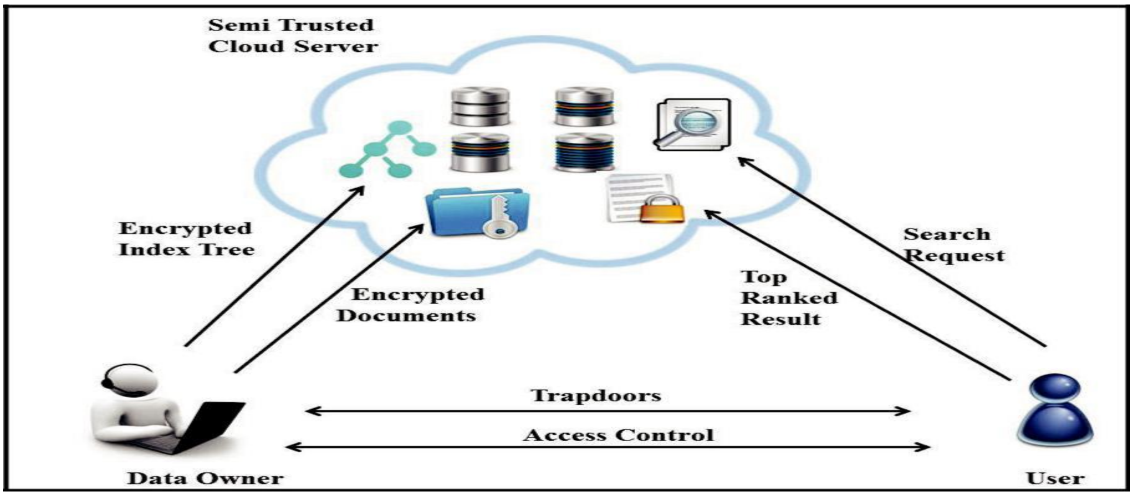

Fig. Searchable Encryption Technique

VI. PROBLEM FORMULATION

- F-> the plaintext document collect a sequence of keywords.
- C -> The encrypted document collection stored in the cloud server, denoted as C= $\{c_1; c_2; \ldots ; c_n\}$.
- T->The unencrypted form of index tree for the whole document collection F .
- I->The searchable encrypted tree index generated from T.
- Q->The query vector for keyword set W_q.
- TD-> The encrypted form of Q, which is named as trapdoor for the search request.
- D_u->The index vector stored in tree node u whose dimension equals to the cardinality of the dictionary
- W.->Note that the node u can be either a leaf node or an internal node of the tree.
- I_{u}->The encrypted form of D_u
- Vector space model and relevance score function. Vector space model along with TF*IDF rule is widely used in plaintext information retrieval, which efficiently supports ranked multi-keyword search [34]. here TF calculated on basis of how many times keywords appear in the query.
- IDF is obtained by dividing the cardinality of the documents collection by the number of document containing the keywords.

VII. MODULES

1. Index Construction of UDMRS Scheme:
Unencrypted dynamic multi keyword ranked search scheme build on basisi of vector space model and KBB tree
There are two secure search scheme:
a.BDMRS
b.EDMRS
In the Index construction first it will generate tree node for each documents, theses nodes are leaf node of index tree.

2. Process of UDMRS Scheme :
The search procedure of the UDMRS scheme is a recursive methodology upon the tree, named as "Greedy Depth first Search " algorithm. We add to an outcome list meant as RList, whose components is described as ⟨RScore; FID⟩. Here, the RScore is the significance score of the archive fFID to the question. The RList stores the k got to reports with the biggest pertinence scores to the inquiry. The rundown's components are positioned in sliding request as indicated by the RScore, and will be upgraded opportune amid the search process.

3.BDMRS Scheme:
In view of the UDMRS scheme, we build the essential element multi-keyword ranked search (BDMRS) scheme by utilizing the secure KNN algorithm [5]. The BDMRS scheme is intended to accomplish privacy preserving in the known cipher text model. BDMRS scheme can secure the IndexConfidentiality and Query Confidentiality in the known cipher text model [6], [7], [8].

4. EDMRS Scheme:
Cloud server has the capacity interface the same search requests by following way of visited nodes. The Cloud server recognize a keyword as the standardized TF distribution of the keyword can be precisely acquired from the last computed relevance scores. A heuristic strategy to further enhance the security is to break such correct quality. Hence, we can acquaint some tunable haphazardness with exasperate the significance score estimation. Likewise, to suit diverse users' inclinations for higher exact positioned results or better protected keyword privacy, the arbitrariness are set movable.

5. Dynamic Update Operation of DMRS:
After insertion or deletion of a record, we require to update synchronously the index. Since the index of DMRS scheme is planned as a balanced binary tree, the dynamic operation is done by redesigning hubs in the list tree. The report on record is just in view of archive recognizes, and no entrance to the substance of records is required.

VIII. CONCLUSIONS

In this paper, we propose secure search scheme supporting multi-keyword ranked search over encrypted cloud data. We make contributions mainly in two aspects: similarity ranked search for more accurate search result with count and tree-based searchable index for more efficient searching. In term of accuracy, we adopt the vector space model combined with cosine measure to evaluate the similarity between search request and document and acquire accurate search result instead of undifferentiated result. For the efficiency aspect, we propose a tree-based index structure. We propose a secure scheme to meet privacy requirements in the threat model. Finally, we analyse the performance of our scheme in detail by the experiment on real-world dataset. But, there still exist some problems, such as dynamic update for searchable index. We will do more research in the future.

REFERENCES

- K. Ren, C. Wang, Q. Wang et al., “Security challenges for the public cloud,” IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.
- S. Kamara and K. Lauter, “Cryptographic cloud storage,” in Financial Cryptography and Data Security. Springer, 2010, pp. 136– 149.
- C. Gentry, “A fully homomorphic encryption scheme,” Ph.D. dissertation, Stanford University, 2009.
- O. Goldreich and R. Ostrovsky, “Software protection and simulation on oblivious rams,” Journal of the ACM (JACM), vol. 43, no. 3, pp. 431–473, 1996.
- D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, “Public key encryption with keyword search,” in Advance in Cryptology Eurocrypt 2004.Springer,2004 pp. 506–522.
- C. Wang, N. Cao, K. Ren, and W. Lou, “Enabling secure and efficient ranked keyword search over outsourced cloud data,” Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8, pp. 1467–1479, 2012.
- N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” in IEEE INFOCOM, April 2011, pp. 829–837.
- 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 Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security. ACM, 2013, pp. 71–82.