Protected And Dynamic Keyword Search
Rank Scheme For Cloud Database

ALMASHPUR PRUTHVI
CSE Department
CMR Technical Campus
Kandlakoya, Medchal Road, Hyderabad, India.

N. BHASKAR
Assoc.Professor, CSE Department
CMR Technical Campus
Kandlakoya, Medchal Road, Hyderabad, India.

Abstract- The major aim of this paper is to solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) at the time of protecting exact method wise privacy in the cloud computing concept. Data holders are encouraged to outsource their difficult data management systems from local sites to the business public cloud for large flexibility and financial savings. However for protecting data privacy, sensitive data have to be encrypted before outsourcing, which performs traditional data utilization based on plaintext keyword search. As a result, allowing an encrypted cloud data search service is of supreme significance. In view of the large number of data users and documents in the cloud, it is essential to permit several keywords in the search demand and return documents in the order of their appropriate to these keywords. Similar mechanism on searchable encryption makes centre on single keyword search or Boolean keyword search, and rarely sort the search results. In the middle of various multi-keyword semantics, deciding the well-organized similarity measure of “coordinate matching,” it means that as many matches as possible, to capture the appropriate data documents to the search query. Particularly, we consider “inner product similarity” i.e., the amount of query keywords shows in a document, to quantitatively estimate such match measure that document to the search query. Through the index construction, every document is connected with a binary vector as a sub index where each bit characterize whether matching keyword is contained in the document. The search query is also illustrates as a binary vector where each bit means whether corresponding keyword appears in this search request, so the matched one could be exactly measured by the inner product of the query vector with the data vector. On the other hand, directly outsourcing the data vector or the query vector will break the index privacy or the search privacy. The vector space model facilitate to offer enough search accuracy, and the DES encryption allow users to occupy in the ranking while the popularity of computing work is done on the server side by process only on cipher text. As a consequence, data leakage can be eradicated and data security is guaranteed.

Keywords— Multi-Keyw ord Ranked Search Over Encrypted Cloud Data; OTP; Product Resemblance; Cloud; Data Owners

I. INTRODUCTION

Cloud Computing is a new but increasingly mature model of enterprise IT infrastructure that provides on-demand high quality applications and services from a shared pool of configuration computing resources. The cloud customers, individuals or enterprises, can outsource their local complex data system into the cloud to avoid the costs of building and maintaining a private storage infrastructure. However, some problems may be caused in this circumstance since the Cloud Service Provider (CSP) possesses full control of the outsourced data. Unauthorized operation on the outsourced data may exist on account of curiosity or profit. To protect the privacy of sensitive information, sensitive data (e.g., emails, photo albums, personal health records, financial records, etc.) should be encrypted by the data owner before outsourcing, which makes the traditional and efficient plaintext keyword search technique useless. The simple and awkward method of downloading all the data and decrypting locally is obviously impractical. So, two aspects should be concentrated on to explore privacy-preserving effective search service. Firstly, ranked search, which can enable data users to find the most relevant information quickly, is a very important issue. The number of documents outsourced to the cloud is so large that the cloud should have the ability to perform search result ranking to meet the demand for effective data retrieval. Secondly, multi-keyword search is also very important to improve search result accuracy as single keyword search often return coarse search results.

In this paper, we propose a practically efficient and flexible searchable encrypted scheme. To address multi-keyword search and result ranking, we use Vector Space Model (VSM) to build document index. To improve search efficiency, we use a tree-based index structure which is a balanced binary tree. We construct the searchable index tree based on the document index vectors. Our encryption scheme can meet the privacy requirements in the threat model.

II. EXISTING SYSTEM

The encrypted data to the cloud and execute keyword search over ciphertext domain. Due to
different cryptography Primitives, searchable encryption schemes can be constructed using public key based cryptography, or symmetric key based cryptography. Song et al. proposed the first symmetric searchable encryption (SSE) scheme, and the search time of their scheme is linear to the size of the data collection. Goh [8] proposed formal security definitions for SSE and designed a scheme based on Bloom filter. The search time of Goh’s scheme is $O(n)$, where $n$ is the cardinality of the document collection. Curtmola et al. [10] proposed two schemes (SSE-1 and SSE-2) which achieve the optimal search time. Their SSE-1 scheme is secure against chosen-keyword attacks (CKA1) and SSE-2 is secure against adaptive chosen-keyword attacks (CKA2). These early works are single keyword boolean search schemes, which are very simple in terms of functionality. Afterward, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity, multi-keyword boolean search, ranked search, and multi-keyword ranked search etc.

Multi-keyword boolean search allows the users to input multiple query keywords to request suitable documents. Among these works, conjunctive keyword search schemes only return the documents that contain all of the query keywords. Disjunctive keyword search schemes return all of the documents that contain a subset of the query keywords. Predicate search schemes are proposed to support both conjunctive and disjunctive search. All these multikeyword search schemes retrieve search results based on the existence of keywords, which cannot provide acceptable result ranking functionality. Ranked search can enable quick search of the most relevant data. Sending back only the top-k most relevant documents can effectively decrease network traffic. Some early works have realized the ranked search using order-preserving techniques, but they are designed only for single keyword search. Cao et al. realized the first privacy-preserving multi-keyword ranked search scheme, in which documents and queries are represented as vectors of dictionary size. With the “coordinate matching”, the documents are ranked according to the number of matched query keywords. However, Cao et al.’s scheme does not consider the importance of the different keywords, and thus is not accurate enough. In addition, the search efficiency of the scheme is linear with the cardinality of document collection.

Sun et al. presented a secure multi-keyword search scheme that supports similarity-based ranking. The authors constructed a searchable index tree based on vector space model and adopted cosine measure together with TF×IDF to provide ranking results. Sun et al.’s search algorithm achieves better-than-linear search efficiency but results in precision loss.

O’reckik et al. proposed a secure multikeyword search method which utilized local sensitive hash (LSH) functions to cluster the similar documents. The LSH algorithm is suitable for similar search but cannot provide exact ranking. In , Zhang et al. proposed a scheme to deal with secure multi-keyword ranked search in a multi-owner model. In this scheme, different data owners use different secret keys to encrypt their documents and keywords while authorized data users can query without knowing keys of these different data owners. The authors proposed an “Additive Order Preserving Function” to retrieve the most relevant search results. However, these works don’t support dynamic operations.

### III. PROPOSED SYSTEM

In the Proposed work, we will discover checking the integrity of the rank order in the search result analysing the cloud server is untrusted. To advise OTP (one Time Password) as our upcoming work. This OTP used to see information in cloud and it can be used once only in a time, when you search a file and be likely to see the file, the OTP will transmit to email and we receive the OTP and apply to see the file.

In this technique the following are the different things which we have to implement

i) Cloud Setup

ii) Cryptography cloud Storage

iii) Vector Model

#### Cloud Setup

Firstly, we have to setup data owner and cloud server. So the data owner will then push the data into the cloud servers. When users outsource their confidential data onto the cloud, the cloud service providers are capable to control and check the data and the communication between users and the cloud will be secured.

#### Cryptography cloud Storage

Secondly, while the data is uploaded into the Estorage and retrieve services. Since data may have confidential information, the cloud servers cannot be fully hand over in protecting data. For this cause, outsourced files must be encrypted. Any kind of information leakage that would change data privacy are regarded as Unacceptable.

#### Vector Model

We used a series of searchable symmetric encryption systems that have been allowing search on cipher text. In the earlier, files are ranked only by the number of get back keywords, which damage search correctness.
IV. CONCLUSION

In this paper, a secure, efficient and dynamic search scheme is proposed, 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. The security of the scheme is protected against two threat models by using the secure kNN algorithm. Experimental results demonstrate the efficiency of our proposed scheme. There are still many challenge problems in symmetric SE schemes. In the proposed scheme, the data owner is responsible for generating updating information and sending them to the cloud server. Thus, the data owner needs to store the unencrypted index tree and the information that are necessary to recalculate the IDF values. Such an active data owner may not be very suitable for the cloud computing model. It could be a meaningful but difficult future work to design a dynamic searchable encryption scheme whose updating operation can be completed by cloud server only, meanwhile reserving the ability to support multi-keyword ranked search. In addition, as the most of works about searchable encryption, our scheme mainly considers the challenge from the cloud server. Actually, there are many secure challenges in a multiuser scheme. Firstly, all the users usually keep the same secure key for trapdoor generation in a symmetric SE scheme. In this case, the revocation of the user is big challenge. If it is needed to revoke a user in this scheme, we need to rebuild the index and distribute the new secure keys to all the authorized users. Secondly, symmetric SE schemes usually assume that all the data users are trustworthy. It is not practical and a dishonest data user will lead to many secure problems. For example, a dishonest data user may search the documents and distribute the decrypted documents to the unauthorized ones. Even more, a dishonest data user may distribute his/her secure keys to the unauthorized ones. In the future works, we will try to improve the SE scheme to handle these challenge problems.

V. REFERENCE


