International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

IJIEMR Transactions, online available on 26th Sept2017. Link :http://www.ijiemr.org/downloads.php?vol=Volume-6&issue=ISSUE-8

Title: MULTIUSER MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED CLOUD

Volume 06, Issue 08, Pages: 416– 423.

Paper Authors

G. RAJANI, M VENKATESH NAIK
CRIT , Anantapur

.
.

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

MULTIUSER MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED CLOUD

G. RAJANI[1], M VENKATESH NAIK [2]
[1]PG Scholar, CSE Department, CRIT , Anantapur
[2] Assistant Professor, CSE Department, CRIT , Anantapur

ABSTRACT: Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and the widely-used TF_IDF model are combined in the index construction and query generation. We construct a special tree-based index structure and propose a "Greedy Depth-first Search" algorithm to provide efficient multi-keyword ranked search. The secure KNN algorithm is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encrypted index and query vectors. In order to resist statistical attacks, phantom terms are added to the index vector for blinding search results. Due to the use of our special tree-based index structure, the proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly. Extensive experiments are conducted to demonstrate the efficiency of the proposed scheme.

Keywords: secure multi-keyword ranked search

1 .INTRODUCTION

Cloud computing is a conversational phrase used to express a variety of dissimilar types of computing ideas that occupy large number of computers that are connected through a real-time communication network i.e Internet. In science, cloud computing is the capability to run a program on many linked computers at the same time. The fame of the term can be recognized to its use in advertising to sell hosted services in the sense of application service provisioning that run client server software on a remote location. Cloud computing relies on sharing of resources to attain consistency and financial system alike to a utility (like the electricity grid) over a network. The cloud also centres on maximize the effectiveness of the shared resources. Cloud resources are typically not only shared by multiple users but as well as dynamically re-allocated as per demand. This can perform for assigning resources to users in dissimilar time zones. For example, a cloud computing service which serves American users during American business timings with a specific application (e.g. email) while the same resources are getting reallocated and serve Indian users during Indian business timings with another application (e.g. web server).

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

This mechanism must take full advantage of the use of computing powers thus decreasing environmental damage as well, since less power, air conditioning and so on, is necessary for the same functions. The expression "moving to cloud" also explains to an organization moving away from a traditional CAPEX model i.e buy the devoted hardware and decrease in value it over a period of time to the OPEX model i.e use a shared cloud infrastructure and pay as you use it. Proponents maintain that cloud computing Permit Corporation to avoid direct infrastructure costs, and focus on projects that distinguish their businesses as an alternative of infrastructure. Proponents also maintains that cloud computing permit schemes to get their applications should run faster, with better manageability and less maintenance, and enable IT to more quickly adjust resources to meet random and changeable business demand.

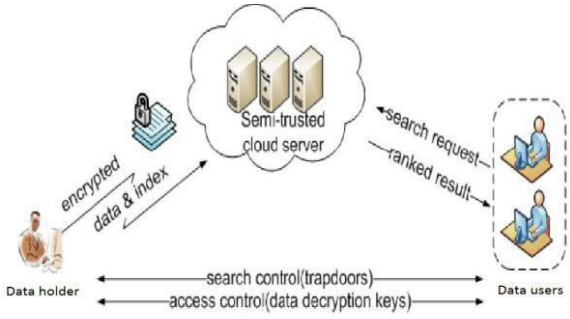

Fig. 1. Architecture of the search over encrypted cloud data.

2. RELATED WORK

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¨ rencik *et al.* proposed a secure multi-keyword 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.

3. MULTI-KEYWORD RANKED SEARCH OVER ENCRYPTED (MRSE) :

Now a day's cloud computing has become essential for many utilities, where cloud customers can slightly store their data into the cloud so as to benefit from on-demand high-quality request and services from a shared pool of configurable computing resources. Its huge suppleness and financial savings are attracting both persons and enterprise to outsource their local complex data management system into the cloud. To safe guard data privacy and struggle unwanted accesses in the cloud and away from, sensitive data, for example, emails, personal health records, photo albums, videos, land documents, financial transactions, and so on, may have to be encrypted by data holder before outsourcing to the business public cloud; on the other hand, obsoletes the traditional data use service based on plaintext keyword search. The insignificant solution of downloading all the information and decrypting nearby is clearly impossible, due to the enormous amount of bandwidth cost in cloud scale systems. Furthermore, apart from eradicating the local storage management, storing data into the cloud supplies no purpose except they can be simply searched and operated. Thus, discovering privacy

preserving and effective search service over encrypted cloud data is one of the supreme importance. In view of the potentially large number of on-demand data users and vast amount of outsourced data documents in the cloud, this difficulty is mostly demanding as it is really difficult to gather the requirements of performance, system usability, and scalability. On the one hand, to congregate the efficient data retrieval requirement, the huge amount of documents orders the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to discover the most appropriate information quickly, rather than burdensomely sorting during every match in the content group. Ranked search can also gracefully remove redundant network traffic by transferring the most relevant data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking operation on the other hand, should not reveal any keyword to related information. To get better the search result exactness as well as to improve the user searching experience, it is also essential for such ranking system to support multiple keywords search, as single keyword search often give up far too common results. As a regular practice specifies by today's web search engines i,e Google search, data users may lean to offer a set of keywords as an alternative of only one as the indicator of their search interest to retrieve the most relevant data. And each keyword in the search demand is able to help narrow down the search result further. "Coordinate matching", as many matches as possible, is an efficient resemblance measure among such multi-keyword semantics to refine the result significance, and has been widely used in the plaintext information retrieval (IR) community. Though, the nature of applying encrypted cloud data search system remains a very demanding task in providing security and maintaining privacy, like the data privacy, the index privacy, the keyword privacy, and many others. Encryption is a helpful method that treats encrypted data as documents and allows a user to securely search through a single keyword and get back documents of interest. On the other hand, direct application of these approaches to the secure large scale cloud data utilization system would not be necessarily suitable, as they are developed as crypto primitives and cannot put up such high service-level needs like system usability, user searching experience, and easy information discovery. Even though some modern plans have been proposed to carry Boolean keyword search as an effort to improve the search flexibility, they are still not sufficient to provide users with satisfactory result ranking functionality. The solution for this problem is to secure ranked search over encrypted data but only for queries consisting of a single keyword. The challenging issue here is how to propose an efficient encrypted data search method that supports multi-keyword semantics without privacy violation. In this paper, we describe and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving exact system wise privacy in the cloud computing concept. Along with various multi-keyword

semantics, select the efficient resemblance measure of "coordinate matching," it means that as various matches as possible, to confine the significance of data documents to the search query. Particularly, inner product similarity the numbers of query keywords show in a document, to quantitatively calculate such similarity assess of that document to the search query. For the period of the index construction, each document is associated with a binary vector as a sub-index where each bit signifies 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 resemblance could be exactly calculated 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 go against the index privacy or the search privacy. To face the challenge of cooperating such multi keyword semantic without privacy breaches, we propose a basic idea for the MRSE using secure inner product computation, which is modified from a secure k-nearest neighbour (kNN) method, and then give two considerably improved MRSE method in a step-by-step way to accomplish different severe privacy needs in two risk models with enlarged attack competence.

4. CONTRIBUTION:

1. We suggest two MRSE schemes based on the
Similarity calculation of "coordinate matching" at the time of assembling different privacy needs in two different threat models.
2. We examine some further improvements of
our ranked search method to maintain more search semantics and dynamic data process.
3. we determine the problem of multi keyword ranked search over encrypted cloud data, and set up a set of privacy needs for such a secure cloud data operation system.
4. Detailed analysis investigating privacy and
Efficiency assurance of the proposed schemes is known, and testing on the real-world data set further show the proposed schemes certainly bring in low overhead on calculation and communication. In this paper we propose two new methods to maintain more search semantics. These methods also study the support of data/index dynamics in the system design.

5. OBJECTIVE OF THE PAPER :

Proposed cloud storage systems that offer privacy, reliability and authentication of client data against a UN trusted cloud provider. This OTP used to see data in cloud and it can be used once only in a time, when you search a file and want to see the file, the OTP will send to the email or to the phone and getting the OTP use the OTP to utilize the file . Presently in the existing system the cloud server hosts third-party data storage and get back services. As information may have sensitive information, the cloud servers cannot be fully hand over in protecting data. For this cause, outsourced files must be encrypted. Any type of data leakage that would involve data privacy is considered as undesirable. To meet the effective data

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

retrieval requirement, the huge amount of documents command the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to find the most appropriate information quickly, rather than burdensomely sorting through every match in the content collection. Ranked search can also gracefully eradicate avoidable network traffic by transferring back only the most appropriate data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking process, yet, should not leak any keyword related information. On the other hand, to progress the search result correctness as well as to improve the user searching experience, it is also essential for such ranking system to maintain multiple keywords search, as single keyword search regularly yields far too common results.

6. 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.

System Architecture

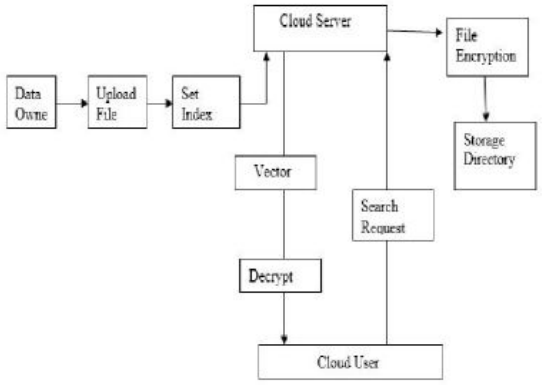

Fig.2 Architecture diagram of the MRSE Implementation.

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.

7. CONCLUSION AND FUTURE WORK

In this paper we describe and solve the problem of multikey word ranked search over encrypted cloud data, and set up a range of privacy requirements. Among various multi-keyword semantics, we select the efficient similarity measure of "coordinate matching," i.e., as many equivalent as possible, to effectively capture the relevance of outsourced documents to the query Keywords, and utilize "inner product similarity" to quantitatively calculate such comparison measure. In order to acquire the test of supporting multi-keyword semantic without privacy violation, we offer a basic idea of MRSE using secure inner product calculation. Then, we give two improved MRSE schemes to attain various severe privacy needs in two different threat models. The further enhancements of our ranked search method, including supporting more search semantics, i.e., TF _ IDF, and dynamic data process. Detailed analyses in investigating privacy and efficiency assurance of proposed schemes are mentioned, and testing on the real-world data set demonstrate our proposed schemes which introduces low transparency on both calculation and communication.

REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou,
"Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM, pp. 829- 837, Apr, 2011.

[2] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M.Lindner, "A Break in the Clouds: Towards a Cloud Definition," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50-55, 2009.

[3] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEE INFOCOM, pp. 693-701, 2012.

*4+ S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.

*5+ A. Singhal, "Modern Information Retrieval: A Brief Overview," IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35- 43, Mar. 2001.

[6] I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Publishing, May 1999.

*7+ D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.

[8] E.-J. Goh, "Secure Indexes," Cryptology ePrint Archive, http:// eprint.iacr.org/2003/216. 2003.

[9] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.

[10] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. 13th ACM Conf. Computer and Comm. Security (CCS '06), 2006.

[11] D. Boneh, G.D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public Key Encryption with Keyword Search," Proc. Int'l Conf. Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2004.

[12] M. Bellare, A. Boldyreva, and A. ONeill, "Deterministic and Efficiently Searchable Encryption," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

[13] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, "Searchable Encryption Revisited: Consistency Properties, sRelation to Anonymous Ibe, and Extensions," J. Cryptology, vol. 21, no. 3, pp. 350-391, 2008.