National Conference on Innovative Trends in Science and Engineering (NC-ITSE'16) ISSN: 2321-8169
Volume: 4 Issue: 7 412 - 416

A Survey on Efficient Privacy Preserving Multi-Keyword Ranked Search for Multiple Data Owners over Encrypted Cloud Storage

Mr. Vaibhav Sawalkar
M.TECH. Student
Dept. of Computer Science Engineering
CIIT Indore, India
Email- vpsawalkar10@gmail.com

Prof. Megha Singh
Assistant Professor & Head
Dept of Computer Science Engineering
CIIT Indore, India
Email-megha_0801@yahoo.co.in

Abstract: Privacy preserving is one of the most important research topics in the data security field. Cloud computing has become an integral part of IT industry, data owners share their outsourced data. Due to these vast amounts of information available on www large number of users attempts to retrieve certain specific data files they are interested in. To eliminate unnecessarily network traffic by not sending back the irrelevant data, ranked keyword search is used. To improve the search result accuracy as well as to enhance the user searching experience, it is necessary for such ranking system to support multi-keyword search is Proposed. The aim of this paper is to study and improve efficiency in privacy preserving over encrypted cloud data using multi-keyword ranked search technique. This paper has reviewed few algorithms and Technique related to multi-keyword ranked search to improve the search result accuracy as well as user searching experience can be enhanced .

Keywords— CLOUD COMPUTING, CSP, PRMSM, MRSE etc.

*****

1 INTRODUCTION

.Search over encrypted data is a technique of great interest in the cloud computing era, because many believe that sensitive data has to be encrypted before outsourcing to the cloud servers in order to ensure user data privacy. Fig 1 Shows Cloud

Computing, in which enterprise's or individual's databases and applications are moved to the servers in the large data centers (i.e. the cloud) managed by the third-party cloud service providers (CSPs) .Today, the latest paradigm to emerge is that of Cloud computing which promises reliable services delivered through next-generation data centers that are built on virtualized compute and storage technologies. Consumers will be able to access applications and data from a "Cloud" anywhere in the world on demand. The consumers are assured that the Cloud infrastructure is very robust and will always be available at any time[15].

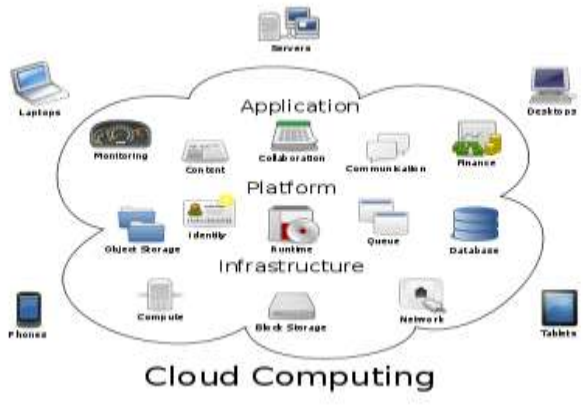


Fig 1 Cloud Computing Architecture

Internet. Privacy preserving is used to preserve the security of fields. If a database has to be shared among several users and some data contained in the database should be prevented by using access control methods in order to guarantee that only authorized people are allowed to have access that sensible information, then the need of privacy and preserving the privacy emerges. Efficient and secure search scheme over encrypted data involves techniques from multiple domains information retrieval for index representation, algorithms for search efficiency, and proper design of cryptographic protocols to ensure the security and privacy of the overall system[3][4]. Cloud computing is the long dreamed vision of computing as a utility, where cloud customers remotely store their data into the cloud so as to enjoy the on-demand high-quality applications and services from a shared pool of configurable computing resources. Its great flexibility and economic savings are motivating both individuals and enterprises to outsource their local complex data management system into the cloud. To protect privacy of data and oppose unsolicited accesses in the cloud and beyond it, sensitive data, for instance, e-mails, personal health records, photo albums, tax documents, and so on, may have to be encrypted by data owners before outsourcing to the commercial public cloud. traditional data utilization service based on plaintext keyword search[6][8].

II LITERATURE SURVEY

Many searching techniques over encrypted cloud data have proposed. S. Deshpande [7] suggested a technique searching over encrypted cloud data using fuzzy keywords. They used Edit distance to quantify keyword similarity and developed two techniques on constructing fuzzy keyword sets to achieve

optimized storage and representation overheads. Cong wang et al. [1] Has proposed a method ranked keyword search over encrypted cloud data using keyword frequency and order preserving encryption. It supports only single keywords at a time. Is the keyword frequency deciding document file score. Rank given to every file based on the relevance score of that file. Top ranked files have sent to users instead all files. To enrich search functionality N. Cao et al. [2] Have proposed a scheme supporting conjunctive keywords search. It is privacy – preserving multi-keyword ranked search technique using symmetric encryption. M. Chou et al. [6] proposed a solution for fuzzy multi-keyword search over encrypted cloud data using privacy aware Bed Tree. They used a co-occurrence probability approach to identify useful multi-keywords for publishing data, documents and relevant fuzzy keyword sets constructed using edit distance. They constructed index tree for all data, documents, where each leaf node having the hash value of a keyword, one or two data vectors that represents n-gram of that keyword and bloom filters for each edit distance value[1][5].

Sr. No.	Paper Title	Objective
1	Secured Multi-keyword Ranked Search over Encrypted Cloud Data	Main focus is on the solution of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving strict system-wise privacy in the cloud computing paradigm.
2	Privacy Preserving Data Sharing With Anonymous ID Assignment	Main objective is to assign user an anonymous ID
3	Providing Privacy Preserving in Cloud Computing	The main idea is protecting individuals' privacy in cloud computing and provides some privacy preserving technologies used in cloud computing services.
4	Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data	This paper has defined and solved the problem of effective yet secure ranked keyword search over encrypted cloud data.
5	Privacy Preserving Keyword Searches on Remote Encrypted Data	Main objective is to to get the access to user's data which is stored remotely from anywhere according to user's convenience
6	Enabling Efficient Fuzzy Keyword Search over Encrypted Data in Cloud Computing	Main idea is to formalize and solve the problem of effective fuzzy keyword search over encrypted cloud data while maintaining keyword privacy.

DESIGN GOAL

To activate ranked search for effective utilization of outsourced cloud data, our system design should simultaneously achieve security and performance guarantees as follows.

1. **Secured Multi-keyword Ranked Search**: To design search schemes which allow multi-keyword query and provide result similarity ranking for valuable data retrieval, instead of returning undifferentiated results.

2**. Privacy**: To prevent cloud server from learning additional information from dataset and index, and to meet privacy requirements.

3. **Effectiveness with high performance:** Above goals on functionality and privacy should be achieved with low communication and computation overhead.

PROBLEM STATEMENT

Data privacy, the data owner can resort to the traditional symmetric key cryptography to encrypt the data before outsourcing, and effectively prevent the cloud server into the outsourced data.

Index privacy, if the cloud server infers any association between keywords and encrypted documents from index. Therefore, the searchable index should be built to prevent the cloud server from acting such kind of association attack.

Keyword Privacy, as users generally wish to have their search from existence showing to others like the cloud server, the most vital concern is to hide what they are searching, i.e., the keywords specified by the corresponding trapdoor. The trapdoor can be generated in a cryptographic way to protect the query keywords.

EXISTING SYSTEM

EXISTING SYSTEM:

Secure search over encrypted data has recently attracted the interest of many researchers. Song et al. first define and solve the problem of secure search over encrypted data. They propose the conception of searchable encryption, which is a cryptographic primitive that enables users to perform a keyword-based search on an encrypted dataset, just as on a plaintext dataset. Searchable encryption is further developed.

Secure search over encrypted cloud data is first defined by Wang et al. and further developed. These researches not only reduce the computation and storage cost for secure keyword search over encrypted cloud data, but also enrich the category of search function, including secure ranked multi-keyword search, fuzzy keyword search, and similarity search.

Drawbacks of Single Keyword Search system*:*

1. Single-keyword search without ranking
2. Boolean- keyword search without ranking
3. Single-keyword search with ranking
4. Do not get relevant data.

THREAT MODEL

In the cloud computing system, the system is not maintained by the data owner and thus it is vulnerable to security threats. We consider an adversary which can intercept the network traffic between data user (or data owner) and the server [18]. We assume the adversary is curious to infer

Additional information from the transmission data (i.e., the encrypted data C, encrypted index I and trapdoor TW′). Based on the information the adversary knows, similar as [6], we consider two threat models for privacy-preserving search in cloud computing system:

Our Propose scheme can handle dynamic data updates efficiently. Our scheme also satisfies the privacy guarantee of searchable encryption schemes as described below.

Known Cipher text Attack: In this threat model, the adversary can intercept the encrypted document collection C, the index I and the query TW′ . In this case, it is computationally intensive for the adversary to conduct the factorization of the polynomial function used in query TW′ and guess the encrypted keywords in H(E(W′)). Thus, the adversary is not able to generate new search request by collecting valid search request. Additionally, the keywords in the query and index are also encrypted: their privacy is also protected as long as the secret key SK = {E(),H()} is kept confidential. Thus, our

multiple keyword search scheme is secure against this threat model[16].

Known Background Attack: In this threat mode, the adversary intends to deduce keywords from the search frequency by using his background information on the dataset. One uniqueness of our multiple keyword search scheme is that it can generate two different query data for the same set of keywords W′ because of the randomly generated dummy keywords. Therefore, the query is not generated in a deterministic manner and the adversary is not able to tell the search frequency

of any keywords. Consequently, the adversary cannot deduce keywords due to the lack of keyword search frequency. Thus, our multiple keyword search scheme is also secure against the known background attack[16].

PROPOSED SYSTEM

In this paper, we propose PRMSM, a privacy preserving ranked multi-keyword search protocol in a multi-owner cloud model.

.

ADVANTAGES

- The proposed scheme allows multi-keyword search over encrypted files which would be encrypted with different keys for different data owners.
- The proposed scheme allows new data owners to enter this system without affecting other data owners or data users, i.e., the scheme supports data owner scalability in a plug-and-play model.
- The proposed scheme ensures that only authenticated data users can perform correct searches. Moreover, once a data user is revoked, he can no longer perform correct searches over the encrypted cloud data.

MOTIVATION

There are some of the motivations for the company's decision of migration to the cloud:

•Standardization:

Standardization means simplifying the system by dealing with less number of configurations, easily facilitated automation and much simpler support. Along with it, the cloud environments being very flexible allows easy provision in various ways. Also it is very user friendly.

• **Virtualization**: Virtualization ensures flexibility,

Increasing the utilization thus being energy efficient. Infrastructure Abstraction and Soft Configuration are characteristics of virtualization.

• **Cost Savings**: Using cloud, capital expenditure can be decreased by not having to buy and maintain costly hardware. A cloud service provider can deploy the data to their high performance systems, with no need to maintain and upgrade expensive software and systems instead the employees can be used to do some productive work for the organization.

• **Better Collaboration**: Good collaboration is the new business success mantra and migration to the cloud makes it much easier to achieve. A more mobile workforce can be achieved who using their own devices can be more productive.

•Improved Network Performance:

If organizations are using remote data centers of their cloud service providers to work on their data, the workload on their on-premise networks can be greatly reduced, thus improving performance of functions using the on premise internal network.

• **Improved Integration and Compatibility**:

The upcoming big data needs of organizations needs them to be capable of accessing and analyzing data stored across

PROPOSED ARCHITECTURE

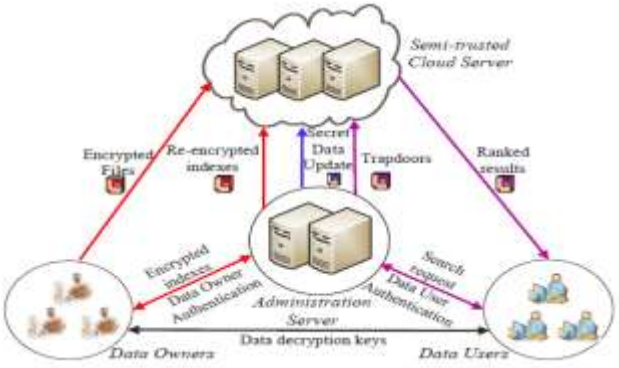
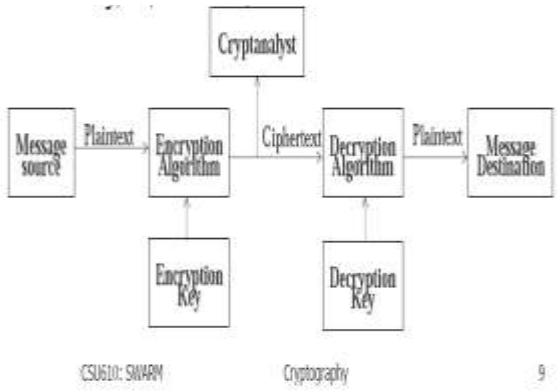


Fig. 2 Architecture of encrypted data search problem

In above Proposed Architecture Consist of *cloud server*, the *data owner*, and the *data user* as shown in Fig.2 The data owner outsources the encrypted dataset and the corresponding secure indexes to the cloud server, where data can be encrypted using any secure encryption technique, such as Advanced Encryption Standard (AES), while the secure index is generated by some particular search-enabled encryption techniques.

ALGORITHMS

EXISTING ALGORITHM



The existing techniques on keyword-based information retrieval, which are widely used on the plaintext data, cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical. All these multi keyword search schemes retrieve search results based on the existence of keywords, which cannot provide acceptable result ranking functionality.

SECURE SEARCH ALGORITHM

According to different data structures, search over encrypted data schemes may use different secure search algorithm to do the match. The inverted index structure allows fast direct intended file retrieval, so the search complexity is constant there.

For example, the indexed keywords can be hashed and then store the associated file list at a table with its address being the hash value . When a user wants to search a keyword of interest, he/she first hashes it and submits the hash value

SIMILARITY-BASED RANKING

To enhance user searching experience and meet more effective data retrieval need, two fundamental aspects have to be considered when designing a practical encrypted data search scheme. On one hand, most of today's search engines on the Internet

(e.g., Google search) allow users to query multiple keywords in one search request instead of only one as the indicator of their search interest. Compared with single keyword query, the main advantage of this multi-keyword search is that it can yield more relevant search results efficiently. On the other hand, ranked search functionality is preferable in the "pay-as-you-go" cloud paradigm. The reason is

that cloud server could conduct relevance ranking operation for data user and return the most relevant set of files, rather than directly sending back the undifferentiated search results to data user. As such, the network traffic between cloud server and data user could be dramatically reduced.

PRIVACY-PRESERVING MULTI-KEYWORD RANKED SEARCH

Multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system. They propose two MRSE schemes based on the similarity measure of coordinate matching while meeting different privacy requirements in two different threat models. One is *known Cipher text model*, where the cloud server is supposed to only know encrypted dataset and searchable index, both of which are outsourced from the data owner.

V CONCLUSION

In this Survey paper we studied various techniques and methods also motivate and problems of supporting efficient ranked keyword search for achieving effective utilization of remotely stored outsourced data in a cloud. We first design the framework definition to provide secure search facility for the sensitive data stored in cloud environment. We also investigate some further enhancements of our ranked search mechanism, including the efficient support of relevance score dynamics.

VI FUTURE ENHANCEMENT

Cloud computing, more and more sensitive data are outsourced to the cloud server to reduce the management cost and enjoy the access services . Need continued research is necessary to further enrich the search functionality and improve the efficiency and scalability of search schemes, another very interesting direction is on virtualization security that tries to secure the execution environment (i.e., virtual machines) in the cloud server.

REFERENCES

[1] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, Mar. 2010.

[2] 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.

[3] S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Se-cure, Scalable, and Fine-Grained Data Access Control in Cloud Computing," Proc.IEEE INFOCOM, 2010.

[4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. KatzA. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, “A view of cloud computing,” Communication of the ACM, Vol. 53, No. 4, pp. 50–58, 2010.

[5] 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, relation to anonymous ibe, and extensions,” J. Cryptol., vol. 21, no. 3, pp. 350–391, 2008.

[6] M.Chuah and W. Hu,"Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data",Distributed Computing Systems Workshops, 2011 31st International Conference,IEEE, (2011).

[7] S. Deshpande, "Fuzzy keyword search over encrypted data in cloud computing",World Journal of Science and Technology, vol. 2, no. 10, (2013).

[8] P. Golle, J. Staddon, and B. Waters, “Secure conjunctive keyword search over encrypted data,” in Proc. of ACNS, 2004, pp. 31–45.

[9] L. Ballard, S. Kamara, and F. Monrose, “Achieving efficient conjunctive keyword searches over encrypted data,” in Proc. of ICICS, 2005.

[10] D. Boneh and B. Waters, “Conjunctive, subset, and range queries on encrypted data,” in Proc. of TCC, 2007, pp. 535–554.

[11] R. Brinkman, “Searching in encrypted data,” in University of Twente, PhD thesis, 2007.

[12] Y. Hwang and P. Lee, “Public key encryption with conjunctive keyword search and its extension to a multi-user system,” in Pairing, 2007.

[13] J. Katz, A. Sahai, and B. Waters, “Predicate encryption supporting disjunctions, polynomial equations, and inner products,” in Proc. Of EUROCRYPT, 2008.

[14] Wenhai Sun et al.,"Verifiable Privacy- Preserving Multi-keyword Text Search in the Cloud Supporting Similarity-based Ranking", Accepted for IEEE Transactions on Parallel and Distributed Systems (TPDS).

[15] Prof Mr. Vijay A Tathe, Prof Ms Deepavali P Patil,"Next Generation Computing on the Internet (GRID) ",International Journal of Scientific and Research Pu blications, Volume 2, Issue 2, February 2012 ISSN 2250-3153

[16] Yanzhi Ren1, Yingying Chen1, Privacy-preserving Ranked Multi-Keyword Search Leveraging Polynomial Function in Cloud Computing”.