International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 3, Special Issue 3 (September '2017)
Proceeding of - International Conference on Recent Trends in Computational Life Science and Information Technology,
Conference held at Tiruppur Kumaran College for Women,Tirupur,Tamilnadu, India.

A STUDY ON SECURITY MECHANISM ON THE ENCRYPTED CLOUD DATA AGAINST THE SEARCH KEYWORDS

V.Muthulakshmi,
M.Phil Research Scholar,
Department of Computer science,
S.N.R SONS College,
Coimbatore, TamilNadu, India.

N.Sumathi,
Professor&Head,
Department of Information Technology,
S.N.R SONS College,
Coimbatore, TamilNadu, India.

Abstract: The large enormous data is outsourced to the cloud for scalable data storage. The outsourced data is encrypted due to privacy and confidentially concerns. However importance of the keyword search on the encrypted data motivates the use of the searchable encryption. In order to encourage the multiple keyword searches, many state of art of approaches in implemented. In this paper, we analyse the multi keyword search mechanisms on the encrypted data with inclusion of data updating. The searchable encryption on multiple keywords is employed using vector space model and Tf-idf concepts to generate the index and query for the multiple random keywords. The encrypted index structure is constructed for data files. During the search operation, the search mechanism integrates the trapdoor of the keywords with index information and returns the matched record to the user. In order to reduce the processing time to the large amount of the data request ranked keyword mechanism can be utilized. The state of approaches is only restricted to the conjunctive searches and it is supported only to exact matches of the keyword. To tackle this implication, we propose a fuzzy based multi keyword search based on LSH function using the hash function to generate index and query. The utilization of this model, misspelled keyword can also be hashed into the query vector. The Euclidean distance is used to capture the keyword similarity. The fuzzy based multi keyword search mechanism can improve both efficiency and accuracy.

Keywords: Cloud Data Storage, Data Encryption, Searchable encryption,Encrypted Index Structure, Ranked Keyword Search

I.INTRODUCTION

Cloud Computing is paradigm which provides large distributed computation capacity and memory space [1]. It enables the data owner and data user to get uninterrupted service through the internet facilities irrespective of time and location across multiple platforms such as personnel computers, laptops and mobile phones. Many Cloud Service providers provide flexible ways to share data over the internet but it suffers from various security threats which stand as primary concern in terms of privacy and confidentiality [2]. The public key cryptosystem gained interest in securing the outsourced data to the cloud. However importance of the keyword search on the encrypted data motivates the use of the searchable encryption. The Searchable encryption has been enhanced in various aspects such as keyword search, Ranked keyword search, multiple keyword Search and finally on multiple Ranked keyword search. The Study of this paper focus on the above said encryption models along the analysis against the data updating in the cloud [3]. The encrypted index structure is build to the extracted keywords from the document. Vector space model and Tf-idf concepts generate the index and query for the multiple random keywords [4]. During the search operation, the search mechanism integrates the trapdoor of the keywords with index information and returns the matched record to the user [5]. The rest of the section is organized as follows, section 2 describes the review of literature followed by section 3 to define the proposed methodology as outline and finally section 4 concludes the study of the paper.

II.REVIEW OF LITERATURES

2.1. Multi keyword search mechanisms on the encrypted data

2.1.1. Personalized Search over Encrypted Outsourced Data

In this literature, searchable encryption scheme over outsourced data has been analysed. However, most encrypted search over outsourced cloud data utilizes the model of "one size fits all" and ignores personalized search intention. Moreover, most of mechanism support only exact keyword search, which greatly affects data usability and user experience. A design a searchable encryption scheme that supports personalized search and improves user search experience remains a very challenging task. Personalized multi-keyword ranked search over encrypted data (PRSE) while preserving privacy in cloud computing has been modelled using semantic ontology Wordnet [6]. A user interest model for is also devised to analyse the user's search history, and adopt a scoring mechanism to express user interest smartly. The limitations of the model of "one size fit all" and keyword exact search are addressed by two PRSE schemes for different search intentions.

2.1.2. Searchable Public-Key Ciphertexts with Hidden Structures for Fast Keyword Search

In this literature, semantically secure public-key searchable encryption schemes take search time linear with the total number of the ciphertext. Hence it makes retrieval from large-scale databases prohibitive. To alleviate this problem, this literature is been analysed in terms of searchable public-key ciphertext with hidden structures (SPCHS) for keyword search as fast as possible without sacrificing semantic security of the encrypted keywords. In SPCHS, all keyword-searchable ciphertext are structured by hidden relations, and with the search trapdoor corresponding to a keyword, the minimum information of the relations is disclosed to a search algorithm as the guidance to find all matching

ciphertextefficiently [7]. A generic SPCHS construction from anonymous identity-based encryption and collision-free full-identity malleable identity-based key encapsulation mechanism (IBKEM) with anonymity is analysed.

2.1.3 Authorized and ranked multi-keyword search over encrypted cloud data

In this literature, symmetric searchable encryption (SSE) technique is analysed in brief. However, search authorization problem is not consider in the particular model as that requires the cloud server only to return the search results to authorized users. An authorized and ranked multi-keyword search scheme (ARMS) over encrypted cloud data by leveraging the ciphertext policy attribute-based encryption (CP-ABE) and SSE techniques is analysed as optimized model to protect the user data [8]. Security analysis demonstrates that the ARMS scheme can achieve confidentiality of documents, trapdoor unlinkability and collusion resistance.

2.1.4. Privacy-Preserving multikeyword Similarity Search over Outsourced Cloud Data

The amount of data generated by individuals and enterprises is rapidly increasing in order to secure and provide reliable access to the data in the cloud as the data could be sensitive, the direct data outsourcing would have the problem of privacy leakage, new privacy preserving multikeyword similarity search model has been analysed. With the emerging cloud computing paradigm, the data and corresponding complex management tasks can be outsourced to the cloud for the management flexibility and cost savings. In particular, with the consideration of the text data, multiple keywords are defined by the data user as query for information retrieval. The cloud returns the files containing more than a threshold number of input keywords or similar keywords, where the similarity here is defined according to the edit distance metric. In order to provide efficiency to data retrieval against the user keyword, blind signature is employed as it provides the user access privacy, along this Bloom filter's bit pattern combined as to speedup of search task at the cloud side. Proposed analysis is to achieve the search is secure against insider threats and efficient in terms of the search time at the cloud side [9].

2.1.5. Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners

Nowadays the Data owners continue to outsource the data to public cloud servers while allowing data users to retrieve this data for their popularity and in order to generate the amount from their publication as information. For privacy concerns, secure searches over encrypted cloud data have motivated the work under single owner and multiple owner models. However, most cloud servers in practice support multiple owners to share the benefits brought by cloud computing. In this literature, a new scheme is analysed to deal with privacy preserving ranked multi-keyword search in a multi-owner model. To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors [10]. Systematically a new secure search protocol also been constructed in parallel. To rank the search results and preserve the privacy of relevance scores between keywords and files, additive order and privacy preserving function family is been enabled. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, dynamic secret key generation protocol and a new data user authentication protocol is fused into the system. Furthermore, proposed model supports efficient data user revocation.

2.1.6. Privacy-preserving top-k keyword similarity search over outsourced cloud data

In this literature, privacy-preserving of top-k keyword similarity search over outsourced cloud data is been analysed. Taking edit distance as a measure of similarity, the similarity keyword sets for all the keywords in the data collection has been build. Then calculate the relevance scores of the elements in the similarity keyword sets by the widely used tf-idf theory are computed. Leveraging the similarity keyword sets and the relevance scores, secure and efficient tree-based index structure for privacy-preserving top-k keyword similarity search has been derived in this literature [11]. To prevent potential statistical attacks, a two-server model is been introduced to separate the association between the index structure and the data collection in cloud servers. Thorough analysis on the validity of search functionality and formal security proofs for the privacy guarantee of solution is depicted.

2.1.7. Analysis of the secure mechanism on outsourced Data in multiple keyword Search

Author	Technique	Advantage	Limitation
Zhangjie Fu ,Kui Ren	Personalized Search over Encrypted Outsourced Data	Includes the personalized Search intension. It increase data usability and user experience	It is provide efficient results for single keyword search
Peng Xu, Qianhong Wu, Wei Wang, Willy Susilo,	Searchable Public-Key Ciphertexts with Hidden Structures for Fast Keyword Search	1. Search time linear with the total number of the ciphertext 2. Find all matching ciphertext efficiently with less information of disclosure.	It is prohibitive for multiple keyword search on secure data
Hongwei Li, Dongxiao Liu, KunJia, Xiaodong Lin	Authorized and ranked multi-keyword search over encrypted cloud data	1. Achieve confidentiality 2.Trapdoor unlinkability 3.collusion resistance.	Single Level Authentication is considered even it may lead to data leakage
Chia-Mu Yu ; Chi-Yuan Chen ; Han-Chieh Chao	Privacy-Preserving Multikeyword Similarity Search	It provide reliable Access to the data and secures the system against the insider attacks	It leads to more cost due to multi keyword data retrieval

Wei Zhang ;Yaping Lin ; Sheng Xiao ; Jie Wu ; Siwang Zhou	Privacy Preserving Ranked Multi-Keyword Search	It supports efficient user revocation	It loses the flexibility in the data manageme nt
TengYi ping ; Cheng Xiang ; Su Sen ; Wang Yulong ; Shuang Kai	Privacy-preserving top-k keyword similarity	Several scores provides for similarity to keywords and data to be retrieved. Improper information retrieval is fully denied	Data Flexibility and cost of the security is increased.

2.2. Importance of Multi-keyword Ranked Search over encrypted cloud data

Multiple keywords ranked search over encrypted cloud data is established in order to protect against the data threats using random keywords. The Method is employed for solving this inaccurate result ranking and sending backs the top-*K* files to the data user, rather than all of the relevant files. This method can dramatically reduce the communication overhead and still meet user's demand. However, such a ranking operation should not leak any other information related to the keywords.

III.OUTLINE OF THE PROPOSED METHODOLOGY

A fuzzy based multi keyword search employs LSH function to utilizing the hash function to generate index and query for both kind of searches. The utilization of this model, misspelled keyword can also be hashed into the query vector. The Euclidean distance is used to capture the keyword similarity. The fuzzy based multi keyword search mechanism can improve both efficiency and accuracy.

IV.CONCLUSION

The Comprehensive study on the multiple ranked Keyword search on the outsourced data in the cloud is presented and it is analysed against the various security measures. The Analysis includes the data updating in the outsourced data in addition to the search process. The Search process is model with encrypted index structure, query generation and keyword index for the data to be retrieved. Moreover this analysis takes into consideration about misspelled words as it works without building the predefined keywords for search process.

V.REFERENCES

[1]. S. Kamara and K. Lauter, "Cryptographic cloud storage," in Proc. Financ. Cryptography Data Secur., 2010, pp. 136–149.
[2]. Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int. Conf. Appl. Cryptography Netw. Secur., 2005, pp. 442–455.
[3]. P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Appl. Cryptography Netw. Secur., 2004, pp. 31–45.
[4]. D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. 4th Conf. Theory Cryptography, 2007, pp. 535–554.
[5]. J. Wang, X. Chen, X. Huang, I. You, and Y. Xiang, "Verifiable auditing for outsourced database in cloud computing," *IEEE Trans. Comput.*, vol. 64, no. 11, pp. 3293–3303, Nov. 2015.
[6]. Zhangjie Fu , Kui Ren , Jiangang Shu , Xingming Sun , Feng Xiao Huang "Enabling Personalized Search over Encrypted Outsourced Data with Efficiency Improvement" in IEEE Transactions on Parallel and Distributed Systems ,Volume: 27, Issue: 9, Sept. 1 2016
[7]. Peng Xu , Qianhong Wu, Wei Wang , Willy Susilo, Josep Domingo-Ferrer, Hai Jin "Generating Searchable Public-Key Ciphertexts With Hidden Structures for Fast Keyword Search" IEEE Transactions on Information Forensics and Security in Volume: 10, Issue: 9, Sept. 2015
[8]. Hongwei Li, Dongxiao Liu, Kun Jia, Xiaodong Lin "Achieving authorized and ranked multi-keyword search over encrypted cloud data" IEEE International Conference onCommunications (ICC), 2015
[9]. Chia-Mu Yu ; Chi-Yuan Chen ; Han-Chieh Chao "Privacy-Preserving Multikeyword Similarity Search Over Outsourced Cloud Data" in IEEE Systems Journal in Volume: 11, Issue: 2, June 2017)
[10]. Wei Zhang ; Yaping Lin ; Sheng Xiao ; Jie Wu ; Siwang Zhou "Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing"in IEEE Transactions on Computers (Volume: 65, Issue: 5, May 1 2016)
[11]. Teng Yiping ; Cheng Xiang ; Su Sen ; Wang Yulong ; Shuang Kai Privacy-preserving top-k keyword similarity search over outsourced cloud data in China Communications (Volume: 12, Issue: 12, December 2015)