INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 6 (2016) PAGES 1295 - 1299
RECEIVED : 29/10/2016. PUBLISHED: 07/11/2016
November 7, 2016

ACHIEVING PRIVACY AND SECURE MULTI-KEYWORD RANKED EXPLORATION OVER ENCRYPTED DATA

[1]S.Nivedhitha, [2]Dr.T.Priyaradhikadevi,
[1]PG Scholar, Dept Of Computer Science Engineering,Mailam Engineering College,Villupuram,
[2]Head Of Department Computer Science Engineering, Mailam Engineering College,Villupuram.

Abstract:
As cloud computing has become prevalent data owners tend to outsource their sensitive information to the cloud. The data have to be encrypted before outsourcing to protect the data privacy. Related works on searchable encryption support either single keyword search or Boolean keyword search which focuses on exact keyword. It does not support minor typos or format inconsistencies. Sorting of results is also rare. In this paper we propose secure fuzzy multi keyword ranked search over encrypted cloud data. This scheme allows multiple keyword in the search request and returns the documents in the order of their relevance to these keywords. A fuzzy keyword set is built from a predefined set of words based on edit distance so that it can return the matching files or closest possible matching files. Here coordinate matching is used based on secure inner product computation to measure similarity. The proposed scheme is secure and privacy preserving while introducing low overhead on computation and communication.

Keywords: Cloud computing, fuzzy keyword search, ranked search, encryption.

1. INTRODUCTION

Capabilities and elastic resource provisioning. Both individuals and enterprises are motivated to outsource their data to the cloud storage server to reduce cost of management. To prevent unauthorized access in the cloud, sensitive data should be encrypted by data owners before outsourcing to the commercial public cloud [11]; thus makes traditional data utilization service based on plaintext keyword search unsuitable for cloud computing. Data encryption will make effective data utilization a very challenging task as there are large numbers of outsourced data files. Downloading all the data and decrypting locally is not practical, as it results in huge amount of bandwidth cost in cloud scale systems. Data encryption also demands the protection of keyword privacy since keywords usually contain important information related to the data files. Thus, exploring privacy preserving and effective search service over encrypted cloud data is important. In cloud computing, data owners share their outsourced data with a number of authorized users. Keyword based retrieval allows users to retrieve files they are interested in. Keyword-based retrieval is widely used in plaintext search schemes, in which user can retrieve relevant files based on the keyword in the search request. However, it is a difficult task in cipher text scenario due to limited operations on encrypted data. The existing searchable encryption techniques allows performing searches securely and effectively but is not suitable in cloud computing scenario as they support only exact keyword search and does not support minor typos and format inconsistencies are not supported. Sometimes users searching input might not exactly match those pre set keywords due to the possible typos, representation inconsistencies and lack of exact knowledge about the data. Simple spell check mechanisms are used to support fuzzy keyword search. However, this approach will not completely solve the problem and sometimes can be ineffective as it requires additional interaction of user to determine the correct word from the candidates generated by the spell check algorithm, which costs extra computation effort for the users. If a user types some other valid keywords by mistake the spell check algorithm will not work because it cannot differentiate between two actual valid words. Due to these new techniques that has

searching flexibility which support both minor typos and format inconsistencies is required. In this paper, we use edit distance to evaluate keywords similarity for the construction of fuzzy keyword sets and a search scheme based on this set. Numerous searchable symmetric encryption schemes have been proposed to enable search on cipher text. However, these traditional schemes enable users to securely retrieve the cipher text, but they support only Boolean keyword search. Schemes presented in [10], [12], [13] show that they support top-k single keyword retrieval under various scenarios. The files should be ranked in the order of relevance by user's interest and only the files with the highest relevance are sent back to users. Ranked search can eliminate unnecessary network traffic by returning only the most relevant data. This ranking operation should not leak any keyword related information for privacy protection. To improve the search result accuracy and to enhance the user searching experience, it is necessary for such ranking system to support multiple keywords search, as single keyword search often yields coarse results. Data users may provide a set of keywords to indicate their search interest in order to retrieve the most relevant data. Each keyword in the search request helps to narrow down the search result further. "Coordinate matching" [14] is an efficient similarity measure to refine the relevance of result, used in the plaintext information retrieval community widely. However, the application of this measure in the encrypted cloud data search system remains a very challenging task because of security and privacy reasons.

2. RELATED WORK

Decrypting it. The first construction of searchable encryption was proposed by Song et al. , where every word in the document is encrypted independently by a special two-layered encryption construction. Goh [8] proposed to construct the indexes for the data files using Bloom filters. For more efficient search, Chang et al. and Curtmola et al. both proposed a single encrypted hash table index to be built for the entire file collection. Traditional single keyword searchable encryption schemes usually build an encrypted searchable index such that its content is hidden to the server unless it is given appropriate trapdoors generated by secret key.

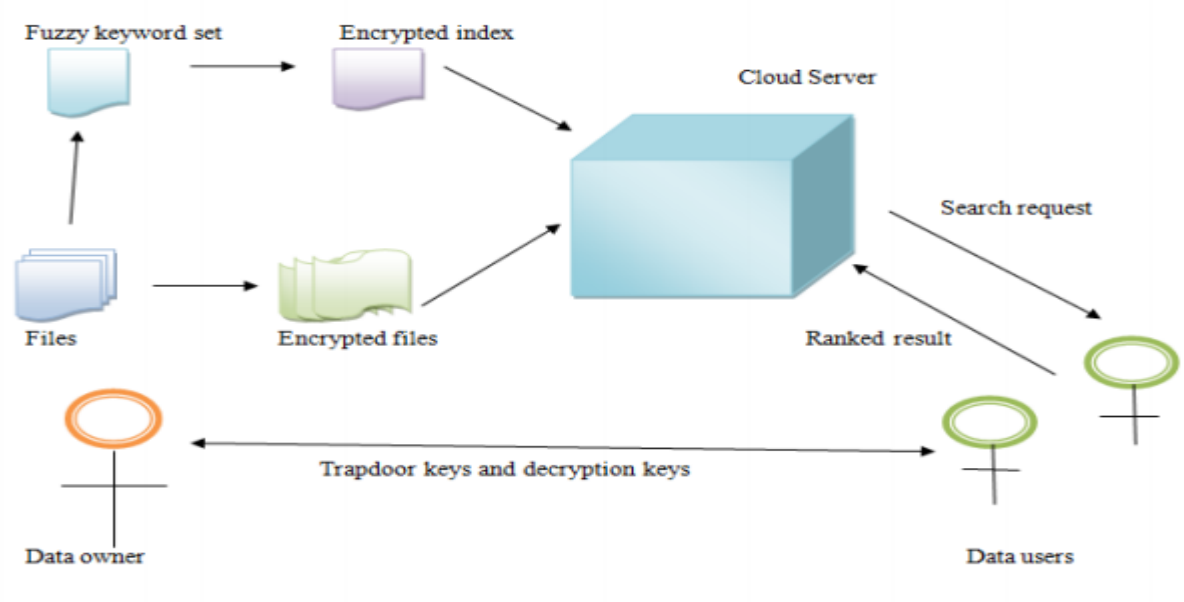

Fig.1. System Architecture

The works utilizes keyword frequency to rank results instead of returning undifferentiated results. Boneh et al. [6] present the first searchable encryption construction, where anyone with public key can write to the data stored on server but only authorized users with private key can search Public key

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 2 ISSUE 6 (2016) PAGES 1295 - 1299
RECEIVED : 29/10/2016. PUBLISHED: 07/11/2016
November 7, 2016

solutions are usually computationally expensive. And the keyword privacy could not be protected in the public key setting since server could encrypt any keyword with public key and then use the received trapdoor to evaluate this cipher text. All these existing schemes support only exact keyword search therefore not suitable for cloud computing, Designs that have been proposed to support Boolean keyword search are still not adequate to provide users with acceptable result ranking functionality. Conjunctive keyword search returns "all-or-nothing," i.e.it only returns those documents where all the keywords specified by the search query appear. Disjunctive keyword search returns only that document that contains a subset of the specific keywords. Predicate encryption schemes are recently proposed to support both conjunctive and disjunctive search. None of existing Boolean keyword searchable encryption schemes support multiple keywords ranked search over encrypted cloud data while preserving.

3. SYSTEM ANALYSIS

Considering a cloud data hosting service involving three different entities, the data owner, the data user along with his ID, and the cloud server. The data owner first registers on cloud using anonymity algorithm for cloud computing services. Before saving user registration information to database present on cloud anonymous algorithm process the data and then anonymous data is saved to registration database. The data owner has a collection of data documents F to be outsourced to the cloud server in the encrypted form C.

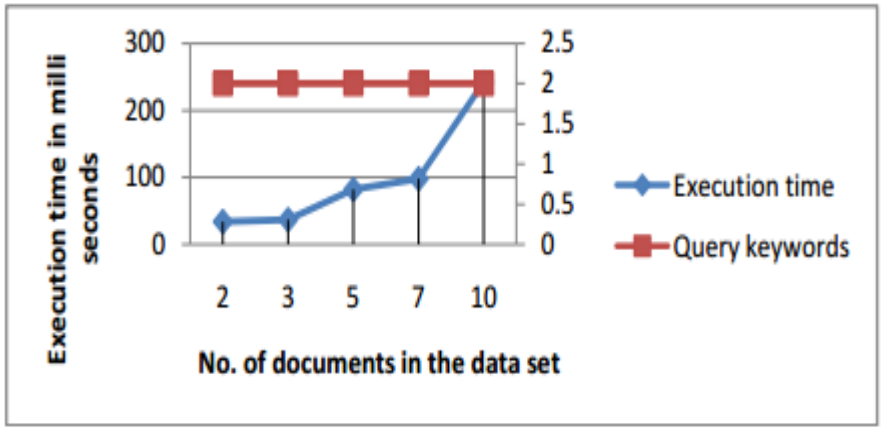

Fig.2. Execution time graph for different number of documents in data set

To enable searching capability over C for effective data utilization, the data owner, will first build an encrypted searchable index I from F before outsourcing , and then outsource both the index I and the encrypted document collection C to the cloud server. The work deals with efficient algorithms for assigning identifiers (IDs) to the users on the cloud in such a way that the IDs are anonymous using a distributed computation with no central authority. Given are N nodes, this assignment is essentially a permutation of the integers {1….N} with each ID being known only to the node to which it is assigned. Our main algorithm is based on a method for anonymously sharing simple data and results in methods for efficient sharing of complex data. To search the document collection for given keywords, an authorized user having an ID acquires a corresponding trapdoor T through search control mechanisms, for example,

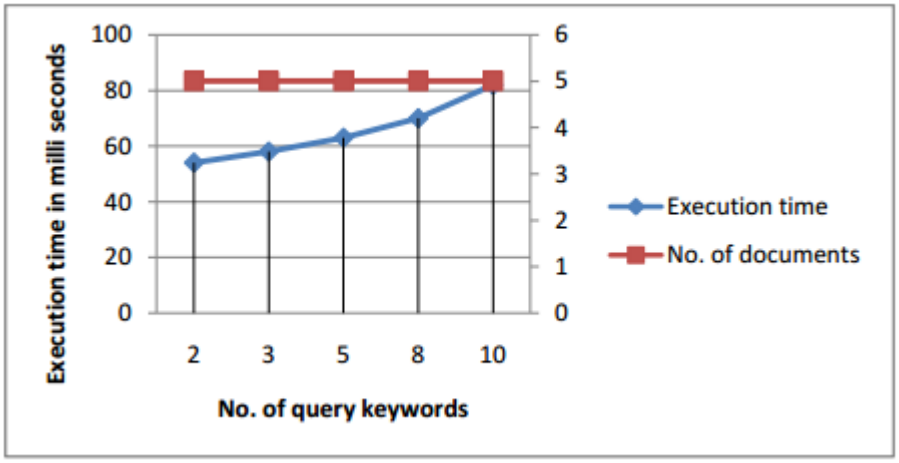

Fig.3. . Execution time graph for different number of query keywords

broadcast encryption. On receiving T from a data user, cloud server is responsible to search the index I and then returns the corresponding set of encrypted documents. In order to improve the document retrieval accuracy, the search result should be ranked by the cloud server according to some ranking criteria (e.g., coordinate matching) and assigning anonymous ID [6] to the user on cloud in order to make the data on cloud more secure. Moreover, to reduce the cost of communication the data user may send an optional number k along with the trapdoor T so that the cloud server only sends back top-k documents that are most relevant to the search query. At last, the access control mechanism is employed in order to manage decryption capabilities given to users and the data collection can be updated in terms of inserting new documents, updating existing ones, and deleting the existing documents.

CONCLUSION

In this paper, we propose the method of secure fuzzy multi keyword ranked search over encrypted cloud data for efficient utilization of remotely stored encrypted cloud data. An advanced technique is designed to construct fuzzy keyword set based on similarity metric of edit distance. Based on the fuzzy keyword set an efficient privacy preserving keyword search scheme is proposed. The efficient similarity measure of "inner product computation" is used to capture the relevance of outsourced documents to the query keywords. Top-k documents are returned to the user based on similarity score, where k is specified by the user. The proposed solution is secure and impose low overhead on computation 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, 2014.

[2] Jiadi Yu, Peng Lu, Yanmin Zhu, Guangtao Xue, Member, IEEE Computer Society, and Minglu Li, "Toward Secure Multikeyword Top k Retrieval over Encrypted Cloud Data", IEEE Transactions, July 2013. [3] Ming Li et al.," Authorized Private Keyword Search over Encrypted Data in Cloud Computing, IEEE proc. International conference on distributed computing systems, June 2011, pages 383-392.

[4] J. Li et al., "Fuzzy Keyword Search Over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM, 2010, pp. 441–45.

[5] Ming Li et al., "Toward Privacy-Assured and Searchable Cloud Data Storage Services", IEEE Transactions on Network, volume 27, Issue 4, July/August 2013.
[6] 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.
[7] Wei Zhou et al., "K-Gram Based Fuzzy Keyword Search over Encrypted Cloud Computing "Journal of Software Engineering and Applications, Scientific Research , Issue 6, Volume 29-32,January2013.
[8] E.-J. Goh, "Secure Indexes," Cryptology ePrint Archive, http://eprint.iacr.org/2003/216. 2003.
[9] M. Li, S. Yu, N. Cao, and W. Lou, "Authorized Private Keyword Search over Encrypted Data in Cloud Computing," Proc. 31st Int'l Conf. Distributed Computing Systems (ICDCS '10), pp. 383-392, June 2011.