International Journal of Research Studies in Computer Science and Engineering (IJRSCSE)
Volume 2, Issue 7, July 2015, PP 7-10
ISSN 2349-4840 (Print) & ISSN 2349-4859 (Online)
www.arcjournals.org

Secured Multi-Keyword Ranked Search over Encrypted Cloud Data

K.Narasimha Reddy, K.Thirupal Reddy[2], Dr.T.Swarna Latha[3]

M.Tech. Student, Dept. of CSE, Narayana Engineering College, Gudur, Andhra Pradesh, India [1]
Assoc Professor, Dept. of CSE, Narayana Engineering College, Gudur, Andhra Pradesh, India [2]
Professor, HOD, Dept. of CSE, Narayana Engineering College, Gudur, Andhra Pradesh, India [3]

Abstract: *Improvement in cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great economic savings and flexibility. But giving better security for data privacy, sensitive data have to be changed to other forms such as encryption before outsourcing, which obsoletes traditional data utilization based on normal keyword search. Thus, enabling an encrypted cloud data search service is of overriding importance. As an enhancement we enhance the existing system we propose an effective-approach to solve the problem of multiple keywords ranked search over encrypted cloud data supporting synonym queries. The main objective of this paper is summarized in two aspects: synonym-based search to support synonym queries and multi-keyword ranked search to achieve more accurate search results. Meanwhile, existing search approaches over encrypted cloud data support only fuzzy keyword or exact search, but not semantics based multi keyword ranked search. Therefore, how to enable an effective searchable system with support of ranked search remains a very challenging problem.*

1. INTRODUCTION

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; this, however, obsoletes the traditional data utilization service based on plaintext keyword search. The insignificant solution of downloading all the data and decrypting locally is clearly impractical, due to the large amount of bandwidth cost in cloud scale systems. Images also contain useful and important information, so proposed system also provides image tagging in MRSE scheme. Moreover, aside from eliminating the local storage management, storing data into the cloud doesn't serve any purpose unless they can be easily searched and utilized. Hence, exploring privacypreserving and effective search service over encrypted cloud data is of great importance. Considering potentially huge number of on-demand data users and large amount of outsourced data documents in the cloud, this problem is particularly challenging as it is extremely difficult to meet also the requirements of performance, system usability, and scalability. Document ranking is provided for fast search, but the priorities of all the data documents is kept same so that the cloud service provider and third party remains unaware of the important documents, thus, maintaining privacy of data. Ranked search can also elegantly eliminate unnecessary network traffic by sending back only the most relevant data, which is highly desirable in the "pay-as-you-use" cloud paradigm. For privacy protection, such ranking operation, however, should not leak any keyword related information. Besides, to improve search result accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keyword search, as single keyword search often yields far too coarse results. As a common practice indicated by today's web search engines (ex. Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most relevant data. Along with the privacy of data and efficient searching schemes, real privacy is obtained only if the user's identity remains hidden from the Cloud Service Provider (CSP) as well as the third party user on the cloud server.

2. Related Work

Organizations, companies store more and more valuable information is on cloud to protect their data from virus, hacking. The benefits of the new computing model include but are not limited to: relief of the trouble for storage administration, data access, and avoidance of high expenditure on hardware mechanism, software, etc. Ranked search improves system usability by normal matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency),As directly outsourcing relevance scores will drips a lot of sensitive information against the keyword privacy, We proposed asymmetric encryption with ranking result of queried data which will give only expected data.

2.1. Existing System

The effective data retrieval need, the large amount of documents demand the cloud server to perform result relevance ranking, instead of returning undifferentiated results. Such ranked search system enables data users to find the most relevant information quickly, rather than burdensomely sorting through every match in the content collection. Ranked search can also elegantly eliminate unnecessary network traffic by sending back only the most relevant data, which is highly desirable in the "pay-as-you-use" cloud paradigm. For privacy protection, such ranking operation, however, should not leak any keyword related information. On the other hand, to improve the search result accuracy as well as to enhance the user searching experience, it is also necessary for such ranking system to support multiple keywords search, as single keyword search often yields far too coarse results.

2.2. Disadvantages of Existing System

The encrypted cloud data search system remains a very challenging task because of inherent security and privacy obstacles, including various strict requirement.

On enrich the search flexibility, they are still not adequate to provide users with acceptable result ranking functionality

3. Proposed System

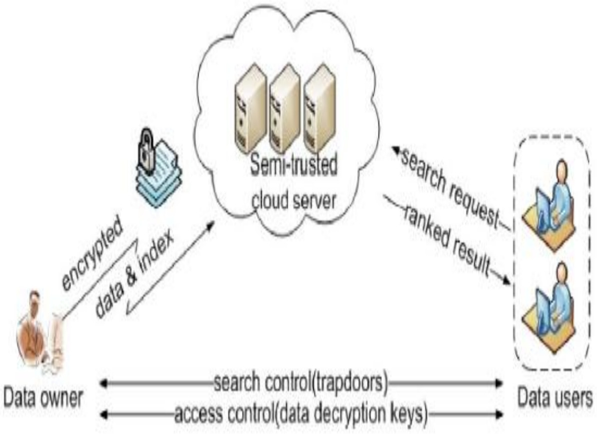
In this paper, for the first time, we define and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving strict system wise privacy in the cloud computing paradigm. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to capture the relevance of data documents to the search query. Specifically, we use "inner product similarity", i.e., the number of query keywords appearing in a document, to quantitatively evaluate such similarity measure of that document to the search query. During the index construction, each document is associated with a binary vector as a sub-index where each bit represents whether corresponding keyword is contained in the document. The search query is also described as a binary vector where each bit means whether corresponding keyword appears in this search request, so the similarity could be exactly measured by the inner product of the query vector with the data vector. However, directly outsourcing the data vector or the query vector will violate the index privacy or the search privacy. To meet the challenge of supporting such multi keyword semantic without privacy breaches, we propose a basic idea for the MRSE using secure inner product computation, which is adapted from a secure k-nearest neighbor (kNN) technique , and then give two significantly improved MRSE schemes in a step-by-step manner to achieve various stringent privacy requirements.

Advantages of Proposed System:

Search result should be ranked by the cloud server according to some ranking criteria.

To reduce the communication cost.

System Architecture:


Semi-trusted cloud server
encrypted data & index
search request
ranked result
Data owner
Data users
search control(trapdoors)
access control(data decryption keys)

The following modules are implemented in this technique

a. Cloud Setup

b. Cryptography cloud Storage

c. Vector Model

Cloud Setup In this module we have setup data owner and cloud server. So the data owner is going push the data into the cloud sever. When users outsource their private data onto the cloud, the cloud service providers are able to control and monitor the data and the communication between users and the cloud will be secured

Cryptography cloud Storage In this module while the data is uploaded into the storage and retrieve services. Since data may contain sensitive information, the cloud servers cannot be fully entrusted in protecting data. For this reason, outsourced files must be encrypted. Any kind of information leakage that would affect data privacy are regarded as unacceptable

Vector Model In this model we used a series of searchable symmetric encryption schemes have been enable search on cipher text. In the former, files are ranked only by the number of retrieved keywords, which impairs search accuracy

4. Conclusion

In this paper, for the first time we define and solve the problem of multi-keyword ranked search over encrypted cloud data, and establish a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, we propose a basic idea of MRSE using secure inner product computation. Then, we give two improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. We also investigate some further enhancements of our ranked search mechanism, including supporting more search semantics, i.e., TF_IDF, and dynamic data operations. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world data set show our proposed schemes introduce low overhead on both computation and communication. In our future work, we will explore checking the integrity of the rank order in the search result assuming the cloud server is untrusted.

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, Relation to Anonymous Ibe, and Extensions,” J. Cryptology, vol. 21, no. 3, pp. 350- 391, 2008.

[14]. 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.