IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

Compound Keyword Level Search to conserve Privacy in access of Encrypted Cloud

Dr. P. Karuppasamy[1], Dr. G. Karthikeyan[2] and Mr. R. Sankarganesh[3]
Professor, Department of Electronics and Communication Engineering[1]
P. S. R Engineering College, Sivakasi, India

Abstract: *With the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. Thus, enabling an encrypted cloud data search service is of paramount importance. Considering the large number of data users and documents in the cloud, it is necessary to allow multiple keywords in the search request and return documents in the order of their relevance to these keywords. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely sort the search results. In this paper, for the first time, we define and solve the challenging problem of privacy preserving multi-keyword ranked search over encrypted cloud data (MRSE).We establish a set of strict privacy requirements for such a secure cloud data utilization system. Among various multikeyword 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. We further use "inner product similarity" to quantitatively evaluate such similarity measure.*

Keywords: Compound Keyword Level.

I. INTRODUCTION

Cloud computing is the long dreamed vision of computing as a utility, where cloud customers can 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 data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data, e.g., emails, personal health records, photo albums, tax documents, financial transactions, etc., 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 trivial solution of downloading all the data and decrypting locally is clearly impractical, due to the huge amount of bandwidth cost in cloud scale systems.

II. OBJECTIVE

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 multikeyword 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.

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

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

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 in two threat models with increased attack capabilities. Our contributions are summarized as follows,

1. For the first time, we explore the problem of multikeyword ranked search over encrypted cloud data, and establish a set of strict privacy requirements for such a secure cloud data utilization system.
2. We propose two MRSE schemes based on the similarity measure of "coordinate matching" while meeting different privacy requirements in two different threat models.

Thorough analysis investigating privacy and efficiency guarantees of the proposed schemes is given, and experiments on the real-world dataset further show the proposed schemes indeed introduce low overhead on computation and communication

III. METHODOLOGY

Considering a cloud data hosting service evolving three different entities, as illustrated in Fig.1 the data owner, the data user, and the cloud server. The data owner has a collection of data documents *F* to be outsourced to the cloud server in the encrypted form *C*. To enable the searching capability over C for effective data utilization, the data owner, before outsourcing, will first build an encrypted searchable index *I* from *F*, and then outsource both the index *I* and the encrypted document collection *C* to the cloud server. To search the document collection for *t* given keywords, an authorized user acquires a corresponding trapdoor *T* through search control mechanisms, e.g., broadcast encryption. Upon receiving *T* from a data user, the cloud server is responsible to search the index *I* and return the corresponding set of encrypted documents. 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, as will be introduced shortly). Moreover, to reduce the communication cost, 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. Finally, the access control mechanism is employed to manage decryption capabilities given to users.

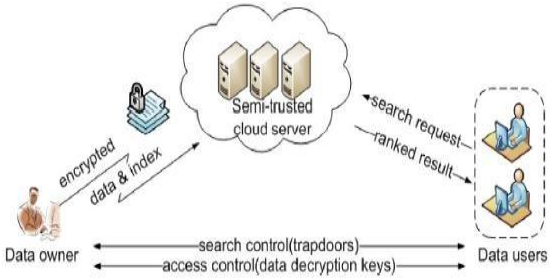

Figure 1: Architecture of the Encrypted Cloud Data

IV. MODULES DESCRIPTION

4.1 Admin Module

This module is used to help the server to view details and upload files with the security. Admin uses the log key to the login time. Before the admin logout, change the log key. The admin can change the password after the login and view the user downloading details and the counting of file request details on flowchart. The admin can upload the file after the conversion of the Zip file format.

4.2 User Module

This module is used to help the client to search the file using the multiple key words concept and get the accurate result list based on the user query. The user is going to select the required file and register the user details and get activation code in mail from the "customerservice404" email before enter the activation code. After user can download the Zip file and extract that file.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

4.3 Owner Data Privacy

Data privacy, the data owner can resort to the traditional symmetric key cryptography to encrypt the data before outsourcing, and successfully prevent the cloud server from prying into the outsourced data. With respect to the index privacy, if the cloud server deduces any association between keywords and encrypted documents from index, it may learn the major subject of a document, even the content of a short document. Therefore, the searchable index should be constructed to prevent the cloud server from performing such kind of association attack. While data and index privacy guarantees are demanded by default in the related literature, various search privacy requirements.

4.4 Keyword Privacy

Keyword privacy as users usually prefer to keep their search from being exposed to others like the cloud server, the most important concern is to hide what they are searching, i.e., the keywords indicated by the corresponding trapdoor. Although the trapdoor can be generated in a cryptographic way to protect the query keywords, the cloud server could do some statistical analysis over the search result to make an estimate. As a kind of statistical information, document frequency (i.e., the number of documents containing the keyword) is sufficient to identify the keyword with high probability in the cloud server knows some background information of the data set; this keyword specific information may be utilized to reverse engineer the keyword.

4.5 Multi-Keyword Search Module

This module is used to help the user to get the accurate result based on the multiple keyword concepts. The users can enter the multiple words query, the server is going to split that query into a single word after search that word file in our database. Finally, display the matched word list from the database and the user gets the file from that list.

V. 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. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset show our proposed schemes introduce low overhead on both computation and communication.

In our future work, we will explore supporting other multikeyword semantics (e.g., weighted query) over encrypted data and checking the integrity of the rank order in the search result.

REFERENCES

[1]. Ning Cao, Member, IEEE, Cong Wang, Member, IEEE, Ming Li, Member, IEEE, KuiRen, Senior Member, IEEE, and Wenjing Lou, Senior Member, IEEE “Privacy- Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data” IEEE Transactions On Parallel And Distributed Systems, VOL. 25, NO. 1, JANUARY2014.

[2]. M. Nabeel and E. Bertino, “Privacy Preserving Delegated Access Control in the Storage as a Service Model,” Proc. IEEE Int’l Conf. Information Reuse and Integration (IRI), 2012.

[3]. E. Bertino and E. Ferrari, “Secure and Selective Dissemination of XML Documents,” ACM Trans. Information and System Security, vol. 5, no. 3, pp. 290-321, 2002.

[4]. G. Miklau and D. Suciu, “Controlling Access to Published Data Using Cryptography,” Proc. 29th Int’l Conf. Very Large Data Bases (VLDB ’03), pp. 898-909,2003.

[5]. N. Shang, M. Nabeel, F. Paci, and E. Bertino, “APrivacy- Preserving Approach to Policy-Based Content Dissemination,” Proc. IEEE 26th Int’l Conf. Data Eng. (ICDE ’10),2010.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

[6]. M.Nabeel, E. Bertino, M. Kantarcioglu, and B.M. Thuraisingham, "Towards Privacy Preserving Access Control in the Cloud," Proc. Seventh Int'l Conf. Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom '11), pp. 172-180,2011.

[7]. M. Nabeel, N. Shang, and E. Bertino, "Privacy Preserving Policy Based Content Sharing in Public Clouds," IEEE Trans. Knowledge and Data Eng., vol. 25, no. 11, pp. 2602- 2614, Nov. 2013.

[8]. S.D.C. di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi and P. Samarati, "Over-Encryption: Management of Access Control Evolution on Outsourced Data," Proc. 33rd Int'l Conf. Very Large Data Bases (VLDB '07), pp. 123-134, 2007.