Secure Ranked Multi-Keyword Search Over Encrypted Cloud Data Using Synonym Queries

[1]A.Arun, [2]J.Danny Anurisha, [3]R.Vignesh, [4]S.Preethi.

PG Scholar
Department of Computer Science and Engineering,
Valliammai Engineering College SRM Nagar
Kattankulathur-603203, Tamilnadu, India.

***Abstract*— Consumer-centric cloud computing phenomenon has been emerging as the development of Smart electronic devices combined with the emerging cloud computing technologies in current years. Different cloud services are provided to the consumer for effective and efficient cloud search service. The consumers rely on most relevant products or data, which is highly desirable in the "pay-as-you use" cloud computing model. As sensitive data (such as photo albums, emails, personal health records, financial records, etc.) are encrypted before outsourcing to cloud, traditional keyword search techniques are useless. Meanwhile, existing search approaches support only exact or fuzzy keyword search, but no semantics-based multi-keyword ranked search. Therefore this paper put forward an effective approach to solve the problem of multi-keyword ranked search over encrypted cloud data supporting synonym queries. The main contribution of this paper is summarized in two aspects: multi-keyword ranked search to achieve more precise search results and synonym-based search to support equivalent queries. The proposed solution is very effective and efficient for multi keyword ranked searching in a cloud environment.**

***Index Terms*— Multikeyword ranked search, Synonym queries, Fuzzy keyword search, CSP-Cloud Service Provider.**

I. INTRODUCTION

In recent years, many consumer electronic devices (e.g. Smartphone) with support of high speed computing combined with the emerging cloud computing paradigm provide a variety of service to the consumers[1]. Consumer-centric cloud computing is a new model of enterprise-level IT infrastructure that provides on-demand high calibre applications and services from a part of configuring computing resources for consumers[3]. On the other hand, some problems may be caused in this context since the Cloud Service Provider (CSP) possesses full control of the outsourced data[5]. There may be unauthenticated operation on the outsourced data on account of curiosity or profit. So sensitive data are encrypted before outsourcing to the cloud. However, encrypted data make the conventional data utilization services based on plaintext keyword search futile[4,6]. The simple and awkward method of downloading all the data and decrypting locally is obviously impractical, because the authenticated cloud consumers must hope to search only specific data rather than all the data[9]. Hence, it is an important thing to explore an effective search service over encrypted outsourced data.

II. RELATED WORK

Current search approaches cannot accommodate searches based on ranking, multi-keywords and semantics-based[7]. To meet the challenge of effective search system, this paper proposes a real time scheme that supports both multi-keyword ranked search and synonym-based search[2]. To address multi-keyword search and result ranking, Vector Space Model (VSM)[10] is used to build document index, that is to say, each document is depicted as a vector where each dimension value is the Term Frequency (TF) weight of its corresponding keyword[8]. The contributions of this paper are summarized as follows: For the first time, a semantics-based multi-keyword ranked search technology over encrypted cloud data which supports synonym queries is proposed. The desired outcomes can be achieved when authenticated cloud customers input the synonyms of the predefined keywords, due to the possible synonym substitution or lack of accurate knowledge about the data. Extensive experiments on the real-world dataset further show the effectiveness and efficiency of proposed solution.

III. SYSTEM DESIGN

The architectural approach shows the secured ranked search in the cloud environment by creating a secure index tree and the documents are encrypted by the data owner. When the data user gives the search request the relevant top k ranked result is displayed. Both the search control (i.e.) encrypted queries and access control (i.e.) data decryption keys are shared only between the authenticated data owner and data user.

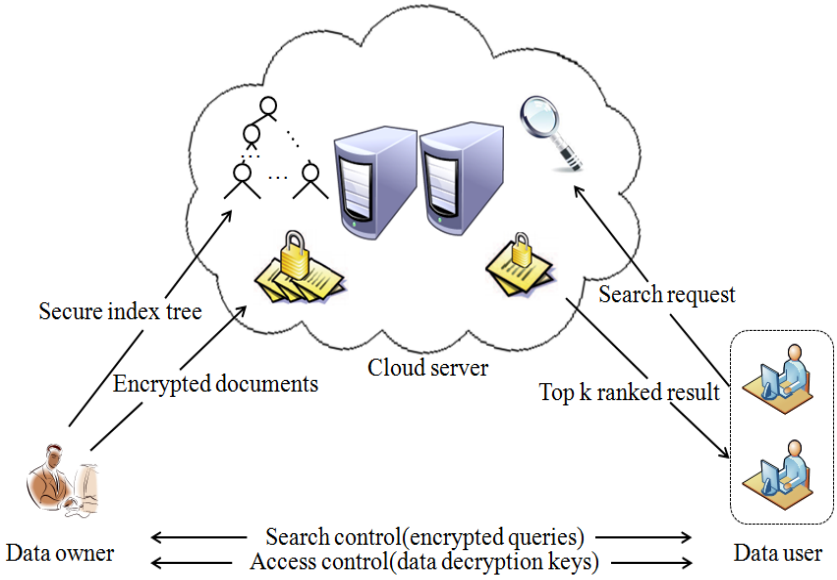

Figure 1: Architectural Approach

IV. CLOUD AUTHORIZATION

The owner registers his/her details to the cloud server provider. Then the registered data owner will upload his/her data to cloud service provider. Upload can be done as private or public. Private files will be uploaded in the data owner private folder and Public files will be uploaded to common public folder. Private files will be accessed only to his/her search but the public files will be accessed to all the data users who accessing that cloud service provider. Here the files will be upload as it normal file without any encryption. Data owner will sometimes act as data user.

V. CONTENT ENCRYPTION AND INDEXING

The data owner will create the index file from the source file which should be uploaded in the cloud service. And then source file will be encrypted with the encrypting algorithm and the encrypted file will be uploaded to the cloud service provider along with its index file. Encryption key is only known to data owner. To create index file we use Text miming Process. The text mining process analyzes the text and also picks literal meaning behind the group of words that constitute the sentence. All the communication to cloud server will be done through web service.

VI. EFFECTIVE CLOUD SEARCH OVER PROCESSED QUERY

The data owner or data user will try give query to the cloud server. The query also will be processed to find the related terms and then it will be searched in the index files. If the terms is present in more than one file, then will search results will be show according to its rank frequency. To which file, frequency is more that file will be listed first. User also will able to give rank to that particular file which will be updated in the cloud server database. Next time that according to frequency and user rank files will be listed to user. If the user selects the particular file then the Encrypted content for the particular file will be provided to the user. Then user has to decrypt that file to access using key which is used for encryption.

VII. OUTCOMES

The proposed system solves the problem of a synonym based multi-keyword ranked search over encrypted cloud data efficiently. It also reduces network traffic and searches most relevant data. Here files are retrieved according to rating or ranking so the frequently visited files are retrieved first so it makes searching process easier.

VIII. CONCLUSION

The proposed system achieves more relevant search results and provides secured index tree and top k ranked result for effective search with authenticated and authorized encryption and decryption facilities in order to upgrade security.

REFERENCES

[1] P.A. Cabarcos, F.A. Mendoza, R.S. Guerrero, A.M. Lopez, and D. DiazSanchez, "SuSSo: seamless and ubiquitous single sign-on for cloud service continuity across devices," IEEE Trans. Consumer Electron.,vol. 58, no. 4,pp. 1425-1433, 2012.

[2] D. Diaz-Sanchez, F. Almenarez, A. Marin, D. Proserpio, and P.A. Cabarcos,"Media cloud: an open cloud computing middleware for content management,"IEEE Trans. Consumer Electron., vol. 57, no. 2, pp. 970-978, 2011
.

[3] S. G. Lee, D. Lee, and S. Lee, "Personalized DTV program recommendation system under a cloud computing environment," IEEE Trans. Consumer Electron., vol. 56, no. 2, pp. 1034-1042, 2010.

[4] 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.

[5] S. Kamara, and K. Lauter, "Cryptographic cloud storage," FC 2010 Workshops, LNCS 6054, PP. 136-149, Jan. 2010.

[6] I. H. Witten, A. Moffat, and T. C. Bell, Managing gigabytes: Compressing and indexing documents and images, Morgan Kaufmann Publishing: San Francisco, May 1999, PP. 36-56.

[7] S. Grzonkowski, and P. M. Corcoran, "Sharing cloud services: user authentication for social enhancement of home networking," IEEE Trans. Consumer Electron., vol. 57, no. 3, pp. 1424-1432, 2011.

[8] R. Sanchez, F. Almenares, P. Arias, D. Diaz-Sanchez, and A. Marin, "Enhancing privacy and dynamic federation in IdM for consumer cloud computing," IEEE Trans. Consumer Electron., vol. 58, no. 1, pp. 95-103,2012.

[9] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," Proceedings of IEEE INFOCOM'10 Mini-Conference, San Diego, CA, USA, pp. 1-5, Mar. 2010.

[10] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," Proceedings of IEEE 30th InternationalConference on Distributed Computing Systems (ICDCS), pp. 253-262, 2010.