International Journal of Pure and Applied Mathematics
Volume 116 No. 9 2017, 225-229
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

ENCRYPTED CLOUD DATA AND EFFICIENT KEYWORD SEARCH

K.P.Kaliyamurthie

Dean, Department of CSE,
Bharath University,Chennai-73.
kpkaliyamurthie@gmail.com

Abstract: Many enterprises are moving their valuable data to cloud due to the rapid expansion of data. With the advantage of "Storage as a service", it will release the burden of data storage and maintenance. It costs less, easily scalable and can be accessed from anywhere any time. With this development, the sensitive information of outsourced data is at risk of unauthorized access. To protect the data privacy, the sensitive data should be encrypted by data owner before outsourcing. To support multi keyword search and result relevance ranking, we adopt "Breadth Deepening Search" algorithm to build the searchable index to achieve accurate search result. Proposed a secure search scheme of RSA algorithm with BSK key generation protocol to meet the privacy requirements.

Keywords: Multi-keyword ranked search over encrypted cloud data, Secret Key, Cloud, Data owners

1. Introduction

Cloud computing is a conversational phrase used to express a variety of dissimilar types of computing ideas that occupy large number of computers that are connected through a real-time communication network i.e. Internet. In science, cloud computing is the capability to run a program on many linked computers at the same time. The fame of the term can be recognized to its use in advertising to sell hosted services in the sense of application service provisioning that run client server software on a remote location. Cloud computing relies on sharing of resources to attain consistency and financial system alike to a utility (like the electricity grid) over a network. The cloud also centers on maximize the effectiveness of the shared resources. Cloud resources are typically not only shared by multiple users but as well as dynamically re-allocated as per demand[1,2,3]. This can perform for assigning resources to users in dissimilar time zones. For example, a cloud computing service which serves American users during American business timings with a specific application (e.g. email) while the same resources are getting reallocated and serve Indian users during Indian business timings with another application (e.g. web server). This mechanism must take full advantage of the use of computing powers thus decreasing environmental damage as well, since less power, air conditioning and so on, is necessary for the same functions[6,7]. The expression "moving to cloud" also explains to an organization moving away from a traditional CAPEX model i.e. buy the devoted hardware and decrease in value it over a period to the OPEX model i.e. use a shared cloud infrastructure and pay as you use it. Proponents maintain that cloud computing Permit Corporation to avoid direct infrastructure costs, and focus on projects that distinguish their businesses as an alternative of infrastructure. Proponents also maintains that cloud computing permit schemes to get their applications should run faster, with better manageability and less maintenance, and enable IT to more quickly adjust resources to meet random and changeable business demand[5].

1.1 Multi-keyword ranked search over encrypted (mrse):

Now a day's cloud computing has become essential for many utilities, where cloud customers can slightly store their data into the cloud to benefit from on-demand high-quality request and services from a shared pool of configurable computing resources. Its huge suppleness and financial savings are attracting both persons and enterprise to outsource their local complex data management system into the cloud. To safe guard data privacy and struggle unwanted accesses in the cloud and away from, sensitive data, for example, emails, personal health records, photo albums, videos, land documents, financial transactions, and so on, may have to be encrypted by data holder before outsourcing to the business public cloud; on the other hand, obsoletes the traditional data use service based on plaintext keyword search. The insignificant solution of downloading all the information and decrypting nearby is clearly impossible, due to the enormous amount of bandwidth cost in cloud scale systems. Furthermore, apart from eradicating the local storage management, storing data into the cloud supplies no purpose except they can be simply searched and operated[8,9,10]. Thus, discovering privacy preserving and effective search service over encrypted cloud data is one of the supreme importance. In view of the potentially large number of

on-demand data users and vast amount of outsourced data documents in the cloud, this difficulty is mostly demanding as it is difficult to gather the requirements of performance, system usability, and scalability.

On the one hand, to congregate the efficient data retrieval requirement, the huge number of documents orders the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to discover the most appropriate information quickly[11,12], rather than burdensomely sorting during every match in the content group. Ranked search can also gracefully remove redundant network traffic by transferring the most relevant data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking operation on the other hand, should not reveal any keyword to related information. To get better the search result exactness as well as to improve the user searching experience, it is also essential for such ranking system to support multiple keywords search, as single keyword search often give up far too common results. As a regular practice specifies by today's web search engines i.e. Google search, data users may lead to offer a set of keywords as an alternative of only one as the indicator of their search interest to retrieve the most relevant data. And each keyword in the search demand can help narrow down the search result further. "Coordinate matching", as many matches as possible, is an efficient resemblance measure among such multi-keyword semantics to refine the result significance, and has been widely used in the plaintext information retrieval (IR) community. Though, the nature of applying encrypted cloud data search system remains a very demanding task in providing security and maintaining privacy, like the data privacy, the index privacy[13,14,15], the keyword privacy, and many others. Encryption is a helpful method that treats encrypted data as documents and allows a user to securely search through a single keyword and get back documents of interest. On the other hand, direct application of these approaches to the secure large scale cloud data utilization system would not be necessarily suitable, as they are developed as crypto primitives and cannot put up such high service-level needs like system usability, user searching experience, and easy information discovery. Even though some modern plans have been proposed to carry Boolean keyword search as an effort to improve the search flexibility, they are still not sufficient to provide users with satisfactory result ranking functionality. The solution for this problem is to secure ranked search over encrypted data but only for queries consisting of a single keyword. The challenging issue here is how to propose an efficient encrypted data search method that supports multi-keyword semantics without privacy violation. In this paper, we describe and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving exact system wise privacy in the cloud computing concept[16,17,18].

Efficiency assurance of the proposed schemes is known, and testing on the real-world data set further show the proposed schemes certainly bring in low overhead on calculation and communication. In this paper, we propose two new methods to maintain more search semantics. These methods also study the support of data/index dynamics in the system design.

2. Contribution

Our contributions are summarized as follows:
1) We design a searchable encryption scheme that supports both the accurate multi-keyword ranked search and flexible dynamic operation on document collection.
2) Due to the special structure of our tree-based index, the search complexity of the proposed scheme is fundamentally kept to logarithmic. And in practice, the proposed scheme can achieve higher search efficiency by executing our "Greedy Depth-first-Search" algorithm. Moreover, parallel search can be flexibly performed to further reduce the time cost
of search process[19].

2.1 Objective of the paper

Cloud storage is a vast model with a mixed network of data users and owners. Data are stored in a pool and it is the responsibility of cloud providers to maintain the data. The Search for any data is one by the data users across the globe and there are various data being hosted in the cloud for any service provider. Objective is to provide an efficient search over the encrypted cloud data by using Breadth Deepening Search algorithm. Secure data by encryption, encryption is best done by the enhanced RSA algorithm. Provide higher search efficiency by executing our "Breadth Deepening Search algorithm" by retrieving the exact and most relevant search. Hacker's unauthorized access are noticed and their access is revoked for security concerns.

3. Proposed system

In this project, Efficient Keyword Search Over Encrypted Cloud Data, which supports ranking and illegal intrusion detection is introduced. The proposed search scheme can achieve deletion and insertion of documents. An efficient search over the encrypted cloud data is achieved using Breadth Deepening Search algorithm(BDS) and the encryption is best done by the enhanced RSA algorithm along with Bit Shift Key(BSK) secures the user data. If there is an

observation of any dishonest activity from the data user, then the access for the user is revoked.

4.System Architecture

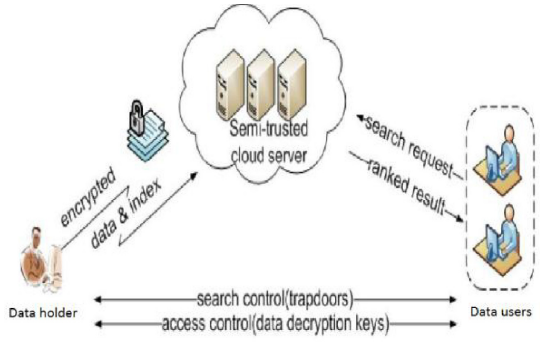

Figure 1. Architecture of the search over encrypted cloud data.

4.1 Search Process of Multi Keyword Search Scheme:

The search process of the UDMRS scheme is a recursive procedure upon the tree, named as "Greedy Depthfirst Search (GDFS)" algorithm. We construct a result list denoted as RList, whose element is defined as RScore; FID. Here, the RScore is the relevance score of the document fFID to the query, which is calculated per Formula (1). The RList stores the k accessed documents with the largest relevance scores to the query. The elements of the list are ranked in descending order
per the RScore, and will be updated timely during the search process. Following are some other
notations, and the GDFS algorithm is described in Algorithm 2.

• RScore(Du;Q) – The function to calculate the relevance score for query vector Q and index vector Du stored in node u, which is defined in Formula (1).
• kthscore – The smallest relevance score in current RList, which is initialized as 0.
• hchild – The child node of a tree node with higher relevance score.

4.2 Algorithm 1 BuildIndexTree(F)

Input: the document collection F = {f1; f2; :::; fn} with the identifiers FID = {FIDjFID = 1; 2; :::; n}.
Output: the index tree T
1: for each document fFID in F do
2: Construct a leaf node u for fFID, with u:ID = GenID(), u:Pl = u:Pr = null, u:FID = FID, and D[i] = TFfFID;wi for i = 1; :::;m;—
3: Insert u to CurrentNodeSet;
4: end for
5: while the number of nodes in CurrentNodeSet is larger than 1 do
6: if the number of nodes in CurrentNodeSet is even, i.e. 2h then
7: for each pair of nodes u′ and u″ in CurrentNodeSet do
8: Generate a parent node u for u′ and u″, with u:ID = GenID(), u:Pl = u′, u:Pr = u″, u:FID = 0 and D[i] = max{u′:D[i]; u″:D[i]} for each i = 1; :::;m;
9: Insert u to TempNodeSet;
10: end for
11: else
12: for each pair of nodes u′ and u″ of the former (2h − 2) nodes in CurrentNodeSet do
13: Generate a parent node u for u′ and u″;
14: Insert u to TempNodeSet;
15: end for
16: Create a parent node u1 for the (2h − 1)-th and 2h-th node, and then create a parent node u for u1 and the (2h + 1)-th node;
17: Insert u to TempNodeSet;
18: end if
19: Replace CurrentNodeSet with TempNodeSet and then clear TempNodeSet;
20: end while
21: return the only node left in CurrentNodeSet, namely,
the root of index tree T ;

4.3 Algorithm 2 GDFS (IndexTreeNode u)

1: if the node u is not a leaf node then
2: if RScore(Du;Q) >kthscore then
3: GDFS(u:hchild);
4: GDFS(u:lchild);
5: else
6: return
7: end if
8: else
9: if RScore(Du;Q) >kthscore then
10: Delete the element with the smallest relevance score from RList;
11: Insert a new element ⟨RScore(Du;Q); u:FID⟩ and sort all the elements of RList;
12: end if
13: return
14: end if

• lchild – The child node of a tree node with lower relevance score.

Since the possible largest relevance score of documents rooted by the node u can be predicted, only a part of the nodes in the tree are accessed during the search process. Fig. 3 shows an example of search process with the document collection F = {fili = 1; :::; 6}, cardinality of the dictionary m = 4, and query vector Q = (0; 0:92; 0; 0:38).

5. Conclusion and future work

In this project, a secure, efficient and dynamic search scheme is proposed, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents. We construct a special keyword balanced binary tree as the index, and propose a Breadth Deepening Search algorithm to obtain better efficiency. The security of the scheme is protected against hacker's user's revocation using BSK algorithm. Experimental results demonstrate the efficiency of our proposed scheme.

In future, the approach to compress the files and to store in the cloud to reduce memory consumption in the cloud and faster download of the compressed files will be analyzed to handle memory challenges. Hence the performance of the search can be improved by avoiding repetition of search.

References

[1] Zhihua Xia, Member, IEEE, Xinhui Wang, Xingming Sun, Senior Member, IEEE,and Qian Wang, Member, IEEE,"A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data"

[2] Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, 2014.

[3] Orencik, M. Kantarcioglu, and E. Savas, "A practical and secure multi-keyword search method over encrypted cloud data," in Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013, pp. 390–397.

[4]Udayakumar R., Kaliyamurthie K.P., Khanaa, Thooyamani K.P., Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, v-29, i-14, pp-86-90, 2014.

[5]Kaliyamurthie K.P., Parameswari D., Udayakumar R., QOS aware privacy preserving location monitoring in wireless sensor network, Indian Journal of Science and Technology, v-6, i-SUPPL5, pp-4648-4652, 2013.

[6]Brintha Rajakumari S., Nalini C., An efficient cost model for data storage with horizontal layout in the cloud, Indian Journal of Science and Technology, v-7, i-, pp-45-46, 2014.

[7]Brintha Rajakumari S., Nalini C., An efficient data mining dataset preparation using aggregation in relational database, Indian Journal of Science and Technology, v-7, i-, pp-44-46, 2014.

[8]Khanna V., Mohanta K., Saravanan T., Recovery of link quality degradation in wireless mesh networks, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4837-4843, 2013.

[9]Khanaa V., Thooyamani K.P., Udayakumar R., A secure and efficient authentication system for distributed wireless sensor network, World Applied Sciences Journal, v-29, i-14, pp-304-308, 2014.

[10]Udayakumar R., Khanaa V., Saravanan T., Saritha G., Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction, Middle - East Journal of Scientific Research, v-16, i-12, pp-1781-1785, 2013.

[11]Khanaa V., Mohanta K., Saravanan. T., Performance analysis of FTTH using GEPON in direct and external modulation, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4848-4852, 2013.

[12]Kaliyamurthie K.P., Udayakumar R., Parameswari D., Mugunthan S.N., Highly secured online voting system over network, Indian Journal of Science and Technology, v-6, i-SUPPL.6, pp-4831-4836, 2013.

[13]Thooyamani K.P., Khanaa V., Udayakumar R., Efficiently measuring denial of service attacks using appropriate metrics, Middle - East Journal of Scientific Research, v-20, i-12, pp-2464-2470, 2014.

[14]R.Kalaiprasath, R.Elankavi, Dr.R.Udayakumar, Cloud Information Accountability (Cia) Framework Ensuring Accountability Of Data In Cloud And Security In End To End Process In Cloud Terminology, International Journal Of Civil Engineering And Technology (Ijciet)Volume 8, Issue 4, Pp. 376–385, April 2017.

[15]R.Elankavi, R.Kalaiprasath, Dr.R.Udayakumar, A fast clustering algorithm for high-dimensional data, International Journal Of Civil Engineering And Technology (Ijciet), Volume 8, Issue 5, Pp. 1220–1227, May 2017.

[16]R. Kalaiprasath, R. Elankavi and Dr. R. Udayakumar. Cloud. Security and Compliance - A Semantic Approach in End to End Security, International Journal Of Mechanical Engineering And Technology (Ijmet), Volume 8, Issue 5, pp-987-994, May 2017.

[17]Thooyamani K.P., Khanaa V., Udayakumar R., Virtual instrumentation based process of agriculture by automation, Middle - East Journal of Scientific Research, v-20, i-12, pp-2604-2612, 2014.

[18]Udayakumar R., Thooyamani K.P., Khanaa, Random projection based data perturbation using geometric transformation, World Applied Sciences Journal, v-29, i-14, pp-19-24, 2014.

[19]Udayakumar R., Thooyamani K.P., Khanaa, Deploying site-to-site VPN connectivity: MPLS Vs IPSec, World Applied Sciences Journal, v-29, i-14, pp-6-10, 2014.

.