SECURE SEARCH OVER ENCRYPTED DATA USING MULTI CLOUD

Mrs. Nethravathi H M[1], Kavana M[2], Rajeshwari M[3], Sandhya V[4], Thara A V[5]
[1]Assistant Professor, [2,3,4,5]U G Students
Dept of Computer Science and Engineering, BGS Institute of Technology, BG Nagar, Mandya-571448

Abstract: ***The advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to 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 cloud, it is crucial for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective data retrieval need. Related works on searchable encryption focus on single keyword search or Boolean keyword search, and rarely differentiate 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), and establish a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. Among various multi-keyword semantics, we choose the efficient principle of "coordinate matching", i.e., as many matches as possible, to capture the similarity between search query and data documents, and further use "inner product similarity" to quantitatively formalize such principle for similarity measurement.***
We first propose a basic MRSE scheme using secure inner product computation, and then significantly improve it to meet different privacy requirements in two levels of threat models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset further show proposed schemes indeed introduce low overhead on computation and communication.

I. INTRODUCTION

Cloud computing is an emerging technology which provides the consumers an easy and on-demand access. It delivers its computing services to the user and gives a storage, data base, network etc. Now a days it has become a very trending technology in many IT-sectors, Business (Public/Private) Organizations, Health-cares, Government Organizations etc..Cloud computing allows features like Adaptability, scalability, Security, Any where access(user can access their data from the cloud from wherever they are) ,On-demand access(user can access the data what they want) etc.
Cloud computing offers different cloud deployment models such as Public cloud, Private cloud, Hybrid cloud ,in that most of the organizations prefers to take services from private cloud in order to upload their documents or data securely. But sometimes the data can be easily accessed or authorized by the cloud service providers illegally, Since the data may contain sensitive information the security will be the main concern while uploading the data into the cloud server. Data encryption has been widely using now a days for security reason in which the plain data or raw data is converted into a cypher data which cannot be read by the unauthorized users in in which many mathematical calculations and algorithms are used.Many encryption models have been used to encrypt the data before outsourcing the data to the cloud.Applying these type of algorithmic encryption models to encrypt the data causes high cost. In order to overcome these problems we are proposing a keyword-based search in this paper that helps the user to search their data securely and efficiently. Although a single keyword search is not enough for searching or storing the data this paper will support a secure multi-keyword ranked search over encrypted data which uses searchable index based algorithm.As the name indicates multi-keyword ranked search which is also known as multi-keyword top-k search, it only returns the k documents with the highest search score.This is motivated by the fact that under certain cases, there are a lot of files matching a user's query, but the user is interested in only a certain percentage of matched files.

II. RELATED WORK

Private searching was proposed by Ostrovsky et al.(referred to as the Ostrovsky scheme in this paper), which allows a user to retrieve files of interest from an untrusted server without leaking any information. However, the Ostrovsky scheme has a high computational cost, since it
requires the cloud to process the query (perform homomorphic encryption) on every file in a collection. Otherwise, the cloud will learn that certain files, without processing, are of no interest to the user. It will quickly become a performance bottleneck when the cloud needs to process thousands of queries over a collection of hundreds of thousands of files.
It has been argued that subsequently proposed improvements, also have the same drawback. Commercial clouds follow a pay-as-you-go model, where the customer is billed for different operations such as bandwidth, CPU time, and so on. Solutions that incur excessive computation and communication costs are unacceptable to customers. To make private searching applicable in a cloud environment existing approach designed a cooperate private searching protocol (COPS), where a proxy server, called the aggregation and distribution layer (ADL), is introduced

between the users and the cloud. The ADL deployed inside an organization has two main functionalities: aggregating user queries and distributing search results. Under the ADL, the computation cost incurred on the cloud can be largely reduced, since the cloud only needs to execute a combined query once, no matter how many users are executing queries. Furthermore, the communication cost incurred on the cloud will also be reduced, since files shared by the users need to be returned only once. Most importantly, by using a series of secure functions, COPS can protect user privacy from the ADL, the cloud, and other users.

III. PROPOSED SYSTEM

The proposed system introduces a novel concept, differential query services, to COPS, where the users are allowed to personally decide how many matched files will be returned. The System Architecture explains the complete overview of the project, before the data owner upload the encrypted document from their local site to the cloud and generate the hashcode for each and every keyword with indexing number on the other hand the data user request for service from the cloud, the user enters the keyword that has to be fetched from the cloud, he will obtain the ranked result, the hashcode will be generated for searched keyword and comparison of that hashcode takes place, the user can access the data only if the hashcode matches. Here we achieve both efficiency and security. If Alice wants to retrieve 2 percent of the files that contain keywords ''A, B'', and Bob wants to retrieve 20 per cent of the files that contain keywords ''A, C''. The cloud holds 1,000 files, where $\{F_1; \ldots; F_{500}\}$ and $\{F_{501}; \ldots; F_{1000}\}$ are described by keywords ''A, B'' and ''A,C'', respectively. In the Ostrovsky scheme, the cloud will have to return 2,000 files. In the COPS scheme, the cloud will have to return 1,000 files. In our scheme, the cloud only needs to return 200 files. Therefore, by allowing the users to retrieve matched files on demand, the bandwidth consumed in the cloud can be largely reduced. The diagram below shows the basic architecture of how the system works.

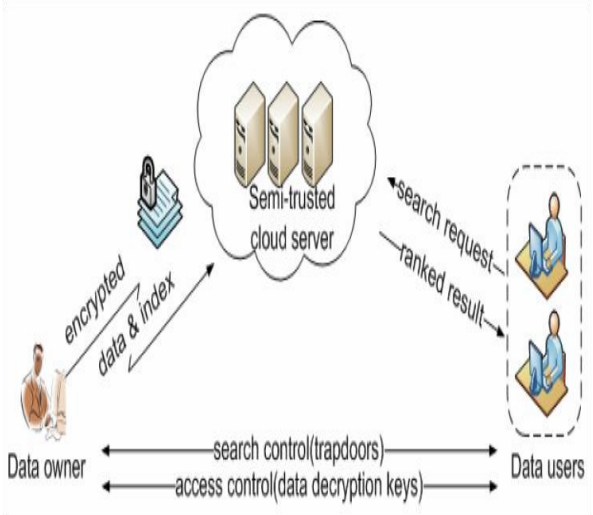

Fig 1. Architecture

Analysing these issues a new scheme is proposed, termed Efficient Information retrieval for Ranked Query (EIRQ), in which each user can choose the rank of his query to determine the percentage of matched files to be returned. The basic idea of EIRQ is to construct a privacy-preserving mask matrix that allows the cloud to filter out a certain percentage of matched files before returning to the ADL. This is not a trivial work, since the cloud needs to correctly filter out files according to the rank of queries without knowing anything about user privacy.

IV. ANALISYS

A. Computing encrypted data

Wouldn't it be nice to be able to…

- Encrypt my data before sending to the cloud
- While still allowing the cloud to search/sort/edit/… this data on my behalf
- Keeping the data in the cloud in encrypted form without needing to ship it back and forth to be decrypted.

B. Algorithm used

Homomorphic encryption

$H = \{\text{KeyGen, Enc, Dec, Eval}\} \leftarrow \text{Eval}_{pk}(f, c)$

$\text{Homomorphic: Dec}_{sk}(\text{Eval}_{pk}(f, \text{Enc}_{pk}(x))) = f(x)$

c^* may not look like a "fresh" ciphertext

As long as it decrypts to $f(x)$

Compact :Decrypting c^*easierthancomputing$f$

$\text{OtherwisewecouldusEval}_{pk}(f, c) = (f, c)$and

$\text{Dec}_{sk}(f, c) = f(\text{Dec}_{sk}(c))$

$|c^*|$independent of the complexity of f

V. RESULTS

We have successfully achieved the result of applying multi-keyword ranked search using homomorphic encryption algorithm. This will give a secure and efficient search over encrypted data.

VI. CONCLUSION

Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval.Secure multi-keyword ranked search scheme over encrypted cloud data, which efficiently search encrypted file from the cloud server.

REFERENCES

[1] Keyword and search engines statistics. http://www.keyworddiscovery.co m/keyword-stats-html?date=2013-01-01(20 13).

[2] Atallah, M.J., Frikken, and K.B:Securely outsourcing linear algebra computations. In: Proceedings of the 5th ACM Symposium on

Information, Computer and Communications Security, pp.48-59,ACM(2010).

[3] Atallah, M.J.,Li, and J.:Secure outsourcing of sequence comparisons. International Journal of Information Security.

[4] 277-287(2005) 4.Attrapadung,N.,Libert, and B.:Functional Encryption for inner product:Achieving constant-size cipher texts with ;adaptive security or support for negation. In:Public Key Cryptography-PKC 2010,pp.384-402.Springer(2010)

[5] Azab, A.M., Ning, P., Zhang, X.:Sice: a hardware-level stronglyisolated computing environment for x86 multi-core platforms. In:Proceedings of the 18th ACM conference on Computer and communications security, pp..375-338.ACM(2011)

[6] Bao, F., Deng., R.H., Ding, X.,Yang, Y.: Private query on encrypted data in multi-user settings. In: Information Security Practice and Experience, pp. 71-85. Springer(2008).

