© 2019 IJSRST | Volume 6 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

Capability of Multi Keyword investigation in Cloud Computing

K. Gayathri, Y. Supriya
G. Pullaiah College of Engineering and Technology, Department of CSE, Kurnool, Andhra Pradesh, India

ABSTRACT

Cloud computing is enormous technical development of this modern era which offers variety of services to satisfy the needs of multiple users. The Cloud service providers charge depending on the user's usage. Imposing confidentiality and scalability on cloud data increases the complexity of cloud computing. Cloud technology has various advantages such as high availability, storage, fast data retrieval, it still has a limitation to overcome which is known as security as sensitive information is centralized into the cloud, and this information must be encrypted and uploaded to cloud for the data privacy and efficient data utilization. As the data becomes complex and number of users are increasing searching of the files must be allowed through multiple keyword of the end users interest. The traditional searchable encryption schemes allows users to search in the encrypted cloud data through keywords, which support only search, i.e., whether a keyword exists in a file or not, without any relevance of data files and the queried keyword. Searching of data in the cloud using Single keyword ranked search results too coarse output and the data privacy is opposed using server side ranking based on order-preserving encryption. In this paper, an efficient clustering technique is used to retrieve encrypted cloud data for multiple related keywords.

Keywords : Cloud Computing, Attribute-Based Encryption, File Hierarchy Document Retrieval.

I. INTRODUCTION

An ever increasing number of individuals and endeavours are inspired to re-appropriate their nearby archive the executives frameworks to the cloud which is a promising data system (IT) to process the unstable extending of information In spite of the benefits of cloud administrations, releasing the delicate data, for example, individual data, organization money related information and government archives, to people in general is a major danger to the information proprietors. Moreover, to make full utilization of the information on the cloud, the information clients need to get to them adaptable and effectively. An instinctive methodology is scrambling the records first and after that re-appropriating the encoded archives to the cloud.

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 rotation, such ranking operation, however, should not leak any keyword related information. On the other hand, to improve search result accuracy as well as to enhance the user searching experience, it is also crucial for such ranking system to support multiple keywords search, as single keyword search often yields far too coarse result. As a common practice indicated by today's web search engines (e.g., Google search), 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. Searchable encryption schemes usually build up an index for

each keyword of interest and associate the index with the files that contain the keyword. By integrating the trapdoors of keywords within the index information, effective keyword search can be realized while both file content and keyword privacy are well-preserved. Although allowing for performing searches securely and effectively, the existing searchable encryption. Techniques do not suit for cloud computing scenario since they support only exact keyword search. The aim of this paper is to achieve an efficient system where any authorized user can perform a search on a remote database with multiple keywords, without revealing neither the keywords he searches for nor the contents of the documents he retrieves.

II. RELATED WORK

Traditional searchable encryption has been widely studied as a cryptographic primitive, with a focus on security definition formalizations and efficiency improvements. So they first introduced the notion of searchable encryption. They proposed a scheme in the symmetric key setting, where each word in the file is encrypted independently under a special two-layered encryption construction. To further enhance search efficiency, a per-keyword-based approach was proposed, where a single encrypted hash table index is built for the entire file collection, with each entry consisting of the trapdoor of a keyword and an encrypted set of related file identifiers. Searchable encryption has also been considered in the public-key setting. Then the first public-key-based searchable encryption scheme construction, with the public key can write to the data stored on the server but only authorized users with the private key can search. As an attempt to enrich query predicates, conjunctive keyword search over encrypted data.

These include the following (a) Secure searchable encryption scheme does not perform any functions when new updates in files or when any modifications are performed. (b) The relevance score algorithm is not updated frequently when there are some modifications in the owner files.

2.1 Contributions: In Cloud Computing, an outsourced file collection might not only be accessed but also updated frequently for various application purposes. Hence, supporting the score dynamics in the searchable index for a secure storage engine which is reflected from the corresponding file collection updates, is thus of practical importance. In our system, we consider score dynamics as adding newly encrypted scores for newly created files, or modifying old encrypted scores for modification of existing files in the file collection. Symmetric key encryption doesn't have major scope in security perspective that's why we are opting MD5 encryption algorithm which is bit more complex, when compared to the traditional algorithms in storing the data. B-Tree indexing and storing of data provides a peak level performance in searching times.

III. EXISTING SYSTEM

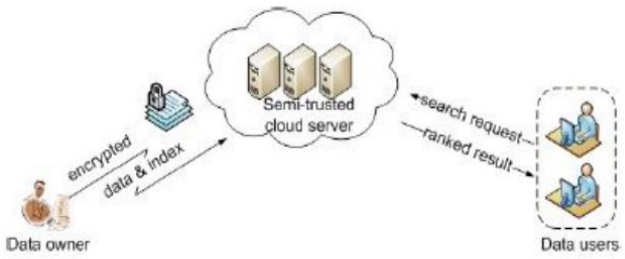

Figure 1. Architecture for search over encrypted cloud data

Design Golas :

To enable ranked searchable symmetric encryption for effective utilization of outsourced and encrypted cloud data under the aforementioned model, our system design should achieve the following security and performance guarantee. Specifically, we have the following goals:

1) Ranked keyword search: to explore different mechanisms for designing effective ranked search

schemes based on the existing searchable encryption framework;
2) Security guarantees: to prevent the clouding server from learning the plaintext of either the data files or the searched keywords, and achieve the "as-strong-as possible" security strength compared to existing searchable encryption schemes;
3) Efficiency: above goals should be achieved with minimum communication and computation overhead.

Disadvantages:

The secure searchable encryption scheme does not perform any function when new updates in files or when any modifications are performed. The relevance score algorithm is not updated frequently when there are some modifications in the owner files.

Proposed System:

In this paper, we solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored encrypted data in Cloud Computing. This is done by developing an efficient clustering algorithm to group the 'related' keywords together. One-to-many order preserving technique protects the score information.

3.1 Overall description: The scenario of the score dynamics mechanism is based on the one-to-one order preserving mapping. An efficient clustering algorithm is used to retrieve encrypted cloud data for multiple related keywords. The multiple related keywords are clustered together and ranked, the information is stored in the index which results in accurate search result when the user searches database with multiple related keywords in the same transaction. The proposed system also ranks cloud data based on end user feedback on top of existing ranking algorithms (which relies on keyword occurrence increases the accuracy of data retrieved).

3.2 Authentication function : Authentication function describes the interface between the user and system and the admin provided the type of authentication. The user is allowed to create his testimonial to login into the system. An admin needs to approve the users created and login approval the user will be allowed to access the application. Authentication is provided by encrypting the user name and password; this protects sensitive information from unauthorized users.

3.3 Clustering algorithm: Clustering is an important application area for many fields including data mining, statistical data analysis, compression, vector quantization, and other business applications. Clustering has been formulated in various ways in the machine learning, pattern recognition, optimization and statistics literature. The fundamental clustering problem is grouping together (clustering) similar data items. During the search process, the user has always desired to input multiple related keywords of his interest rather than a single keyword. Basically any document deal with single concept in brief and the interrelated sub-topics. Grouping the related topics together and forming cluster helps customers to get the desired document of their interest.

Ranked Keyword Search :

Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., Keyword frequency), so achieve the privacy preserving data hosting service in context of cloud computing. Ranked keyword search method protect the relevance score of keyword to leaking the information about keyword for that integrate the new crypto primitive order preserving symmetric encryption and properly modify it for purpose of protect the sensitive weight information.

This technique is providing some functionality. 1. It provides effective protocol, which fulfils the secure ranked search functionality with little relevance score information leakage against keyword privacy. 2. Ranked searchable symmetric encryption scheme is provide as-strong-as-possible security guarantee

compared to previous Searchable symmetric encryption schemes.

The steps of ranked search are shown below.

1. Data owner collects the file and generate the index by extracting the keyword from data files and published index and data files on cloud server.
2. After outsourced the data files user is enable to search and download the data files from cloud server.
3. User can search through only single keyword that is encrypted and using this keyword one trapdoor is generated.
4. Using trapdoor the relevant keyword data files is searched using query and searched data is shown to the user.

Multi Keyword Ranked Search :

In this method searching of cloud data using Privacy Preserving Multi keyword Ranked Search (MRSE). Here basic concept is used is co-ordinate matching. Coordinate matching obtains the similarity between search query and documents. Inner product similarity is also used to describe the multi keyword ranked search over encrypted cloud data (MRSE). The features of this method are, multi-keyword ranked search, privacy preserving, high efficiency is eliminating unnecessary traffic and improve search accuracy.

The steps of ranked search are shown below.

1. Data owner collects the file and generate the index by extracting the keyword from data files and published index and data files on cloud.
2. After outsourced the data files user is enable to search and download the data files from cloud server.
3. User can search through single or multiple keywords that is encrypted and using this keyword one trapdoor is generated.
4. Using trapdoor the relevant keyword data files is searched using query and searched data is shown to the user

IV. CONCLUSION

After the study above two methods are ranked search and multiple keyword ranked search conclude that multi keyword ranked search is better. Multi keyword rank search is enabling semantic keyword search with more accuracy and efficiently because here multiple keywords is used for searching the data files so the frequency of keyword and rank is increased compare to ranked search.

V. REFERENCES

[1]. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia (2009). "Above the clouds: A Berkeley view of cloud computing", University of California, Berkeley, Tech. Rep. UCB-EECS-2009-28.

[2]. S. Kamara and K. Lauter (2010). "Cryptographic cloud storage", in proceeding of financial cryptography: workshop on real-life cryptography protocol. Cloud security Alliance (2009). "Security guidance for critical areas of focus in cloud computing".

[3]. M. Armbrust, April (2010). "A view of cloud computing", Communications of the ACM, vol. 53, no. 4,pp. 50-58.

[4]. C. Wang, N. Cao, K. Ren, and W. Lou, Aug. (2012). "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", Proc. IEEE , Parallel and Distributed Systems.

[5]. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou (2010). "Secure Ranked Keyword Search over Encrypted Cloud Data," Proc. IEEE 30th Int'l

Conf. Distributed Computing Systems (ICDCS '10).

[6]. Ning Caoy, Cong Wangz, Ming Liy, KuiRenz, and WenjingLouy (2013). "Privacy-Preserving Multikeyword Ranked Search over Encrypted Cloud Data" Proc. IEEE Symp. Security and privacy.

Cite this article as :

K. Gayathri, Y. Supriya, "Capability of Multi Keyword investigation in Cloud Computing", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 6 Issue 1, pp. 385-389, January-February 2019.
Journal URL : http://ijsrst.com/IJSRST196144