ISSN: 2277-9655
Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

To Effective Search Of Outsourced Data In Cloud By Using RKS

Deepthi
gopusettideepthi@gmail.com

Abstract

Cloud computing economically enables the paradigm of data service outsourcing. However, to protect data privacy, sensitive cloud data have to be encrypted before outsourced to the commercial public cloud, which makes effective data utilization service a very challenging task. Although traditional searchable encryption techniques allow users to securely search over encrypted data through keywords, they support only Boolean search and are not yet sufficient to meet the effective data utilization need that is inherently demanded by large number of users and huge amount of data files in cloud. In this paper, we define and solve the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy. Specifically, we explore the statistical measure approach, i.e., relevance score, from information retrieval to build a secure searchable index, and develop a one-to-many order-preserving mapping technique to properly protect those sensitive score information. The resulting design is able to facilitate efficient server-side ranking without losing keyword privacy. Thorough analysis shows that our proposed solution enjoys "as strong- as-possible" security guarantee compared to previous searchable encryption schemes, while correctly realizing the goal of ranked keyword search. Extensive experimental results demonstrate the efficiency of the proposed solution.

Keywords: Search in cloud, secured search engines, cloud computing, ranked keyword search,.

Introduction

Cloud Computing enables cloud customers to 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 [1]. The benefits brought by this new computing model include but are not limited to: relief of the burden for storage management, universal data access with independent geographical locations, and avoidance of capital expenditure on hardware, software, and personnel maintenances, etc [2].

Fig1.0 Architecture for search over encrypted cloud data

In above diagram data owner collect the files and will create index file for each of the file and then will encrypt the file using Order Preserving Symmetric Encryption algorithm and then store the data in cloud. In other end user will search for a data in cloud, the searched content will be encrypted format and we need to decrypt the data and then will display the search result in a ranked format. Ranked format of results are obtained using the score calculation of keywords, which is calculated based on document and term frequency algorithm and will be explained in detail in Section III. And also we used concept based search along with this, which greatly improves the efficiency of ranked keyword search. Finally the output by search result will contain relevant data as well as ranking of the word and frequency of the word will be displayed in a ranked format.

Cloud computing is an emerging technology which helps as an utility, through which clients are going to store their data in the cloud server and using applications from a set of computing resources[1]. Here sensitive data is going to be centralized in the server. In some times the cloud server may leaks the data to hackers [2]. The data is going to encrypted before outsourced to achieve privacy. The encryption techniques increase the data utilization from a large amount of data. To retrieve data files we introduced keyword search mechanism. By this mechanism the users are going to retrieve the data files of their interest. In traditional search, encryption techniques the users are

ISSN: 2277-9655
Impact Factor: 1.852

going to search data by using keywords without decrypting it, they support only Boolean keyword search only [2][10]. In cloud computing graded keyword search enhances the system usability by displaying the matching files by the help of relevance score. To achieve security and usability we introduce advanced cryptographic and information retrieval techniques, and using one-to-many order preserving symmetric encryption [3]. Cloud computing offers many benefits, but it also is vulnerable to threats. One of the main threat exist today is the problem of unauthorized users or entities. For avoiding this problem the new technique is developed in this cloud computing is that data owners may share their outsourced data with a large number of users, who might want to only retrieve certain specific data files they are interested in during a given session. One of the most popular ways to do so is through keyword-based search. Such keyword search technique allows users to selectively retrieve files of interest and has been widely applied in plaintext search scenarios.

Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. In cloud computing, the word cloud is used as a metaphor for "the Internet," so the phrase cloud computing means "a type of Internet-based computing," where different services such as servers, storage and applications are delivered to an organization's computers and devices through the Internet. Cloud computing is comparable to grid computing, a type of computing where unused processing cycles of all computers in a network are harnesses to solve problems too intensive for any stand-alone machine**.** This shared IT infrastructure contains large pools of systems that are linked together. Often, virtualization techniques are used to maximize the power of cloud computing.

Framework of Efficient Ranked Keyword and Concept Search

The existing technique resolves the optimization complexities in ranked keyword search and its effective utilization of remotely stored encrypted cloud data. But it limits the further optimizations of the search results by preventing search engine to interact with cloud users to maintain the integrity of actual owner's keyword and the data associated with it.

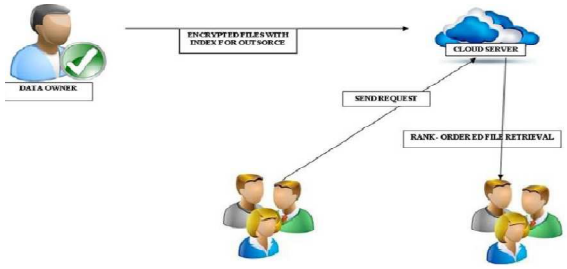

Fig: 2.0 Ranked keyword search in cloud model

Consider an encrypted cloud data hosting service involving three different entities, as illustrated in Fig 2.0 data owner, data user, and cloud server. Data owner has a collection of n data files that he wants to outsource on the cloud server in encrypted form while still keeping the capability to search through them for effective data utilization reasons .Now consider the secure ranked keyword search problem as follows: the search result should be returned according to certain ranked relevance criteria (e.g., keyword frequency-based scores, as will be introduced shortly).The problem with the techniques available for implementing search engine in an environment consists of sensitive outsourced cloud data can be summarized as: a) Lacking of effective mechanisms to ensure the file retrieval accuracy is very difficult. b) Security is not addressed fully and limits search engine's accuracy. The ranked keyword search over encrypted data is to achieve economies of scale for Cloud Computing. This process start from the review of existing searchable symmetric encryption schemes and provides the definitions and framework for this proposed ranked searchable symmetric encryption.

A)Framework Definition :

This process defines and solves the problem of effective yet secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency), thus making one step closer towards practical deployment of privacy-preserving data hosting services in Cloud Computing. It first gives a straightforward yet ideal construction of ranked keyword search searchable symmetric encryption (RSSE)[6] security definition, and demonstrates its inefficiency. To achieve more practical performance, the process then propose a definition for ranked searchable symmetric encryption, and give an efficient design by properly utilizing the existing cryptographic primitive, order-preserving symmetric encryption (OPSE).Relevance Score Calculation[6],

Score (t,Fd) = 1 / |Fd| □ (1+ln fd,t),

ISSN: 2277-9655
Impact Factor: 1.852

where, t- is the term searched by the user, fd,t- denotes the Term Frequency(TF) of the term t in file Fd,
ln- denotes the natural logarithm of TF of the file Fd,
| Fd|- denotes the length of the file.

This process also aim to develop the more efficient in ranked keyword search and provide more security; using the process of TDT4 mechanism and privilege technique. TDT4 mechanism is used for provide the efficient ranked keyword search. In this process information retrieval, a ranking function is used to calculate relevance scores of matching files to a given search request.

To enable ranked keyword search for effective utilization of outsourced cloud data under the aforementioned model, our system design should achieve the following security and performance guarantee.

a) Ranked key word search: For efficient searching process the process use the mechanism of Topic detection and tracking 2004. The search time includes fetching the posting list in the index, decrypting, and rank ordering each entry.
b) Security guarantee: For providing the security in the cloud server, this process uses the privilege method.

b)Mechanisms for Implementation: Topic detection and tracking:

TDT refers to automatic techniques for finding topically related material in streams of data techniques that could be quite valuable in a wide variety of applications where efficient and timely information access is important. For example, a lot of useful information could be gleaned from a multitude of news sources, but no one has the time to watch, listen to, or read carefully each of the many news sources available.

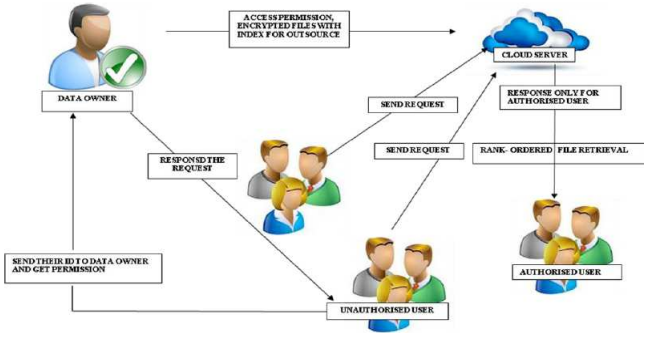

fig:2.1 Ranked keyword search - Secured and efficient

Tasks can vary in focus and size from hypothetical applications to enabling technologies. In brief, the goal of each of the tasks is:
Topic Tracking– detect stories that discuss a target topic,
Link Detection– detect whether a pair of stories discuss the same topic,
Topic Detection– detect clusters of stories that discuss the same topic,
First Story Detection– detect the first story that discusses a topic, and
Story Segmentation– detect story boundaries.

Data Protection

In order to protect the sensitive cloud data in this concept based ranked keyword search we used order preserving symmetric algorithm to encrypt/decrypt the documents. The given below diagram explains the concept of OPS,

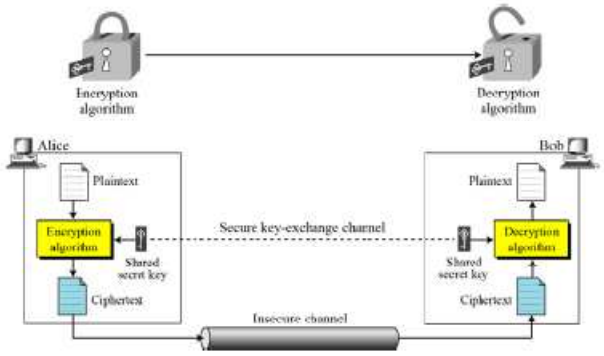

Fig:3.0 Order preserving symmetric Encryption

For example, considered Alice and Bob needs to share the information , in that case if they going to share the plain text format then easily anyone can access or able to view the files.

In order to protect the cloud data we used OPSE which is secure channel of exchanging the data. The above diagram explains the security mechanism that we used, first Alice contains the plain text which will be converted into encryption format and then shared secure key will be generated by OPSE and based on that key Bob can authenticate into application and will decrypt the plaintext which will be in a human readable format.

Conclusion

In this paper, we proposed a searching method to improve the efficiency of ranked keyword search. We gave introduction about the existing searchable encryption framework, it is very inefficient to achieve efficient ranked search. We proposed a efficient one-to-many order preserving mapping function, which allows the effective RSSE to be designed. In additional to that we proposed combination of concept based and keyword based searching techniques .This kind of techniques has the ability to categorize, and search large collections of unstructured text on a conceptual basis. This kind of searching technique is more reliable and efficient search method that is more likely to produce relevant results than traditional searches. Our experimental relevance

score analysis results show that the proposed search methods greatly improve the efficiency of ranked keyword search.

References

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

[2] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," Proc. IEEE 30th Int'l Conf. Distributed Computing Systems (ICDCS '10), 2010.

[3] E.-J. Goh, "Secure Indexes," Technical Report 2003/216, Cryptology ePrint Archive, http://eprint.iacr.org/, 2003..

[4] Dan Boneh,Giovanni Di Crescenzo,Rafail Ostrovsky,Giuseppe Persiano, "Public Key Encryption with Keyword Search" , 2007 [5] Xirong Li , "Harvesting Social Images for Bi-Concept Search" , Proc. IEEE , Multimedia, Aug. 2012 [6] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000. [7] Ming Li, Shucheng Yu†, Ning Cao, Wenjing Lou, "Authorized private keyword searches over encrypted data"