IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

Survey on Privacy Preserving Multi Keyword Search in Cloud Computing

Shreejit Pillai[1], Gaurav Ransing[1], Navanath Ransing[1], Sumit Markad[1], Prof. Nilesh Sable[2]

Bachelor of Engineering, Computer Department, Imperial College of Engineering and Research, Pune, India[1]

Professor, Computer Department, Imperial College of Engineering and Research, Pune, India[2]

Abstract: Due to cloud computing it has become very popular for data owners to outsource data to public cloud servers while allowing data users to retrieve this data. For privacy concerns, a secure search over encrypted cloud data has motivated several research works under the single owner model. However most cloud servers in practice do not serve just one owner instead they support multiple owners to share benefits brought by cloud computing. In this system we propose schemes to deal with privacy preserving ranked multi keyword search in a multi owner environment to enable cloud servers to perform secure search without knowing the actual data of both keywords we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance course between keywords and files, we propose a novel additive order and privacy preserving function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data owners submitting searches we propose a dynamic secret key generation key protocol and a new data user authentication protocol.

Keywords: Cloud computing, multi keyword search, user privacy

I. INTRODUCTION

Cloud computing becomes more and more popular and plays an increasingly important role in our daily lives. It brings users with many benefits such as relief of storage load and flexible data access, which motivate users to store their local data into the cloud. As the cloud services become more prevalent more and more sensitive information such as personal photos, government records and finance data are outsourced into the cloud. To protect the privacy of sensitive data in the cloud the data has to be encrypted by the data owner before outsourcing to the cloud. However the data encryption makes effective data utilization a challenging task when a large amount of data files are present: users may have to download the whole data set from the cloud and then decrypt it to conduct keyword search over the data which is very inefficient when the number of data files is large. Thus an effective keyword searching over encrypted data is of paramount importance, especially the need to provide efficient ranked multiple keyword search which supports a set of input keywords and achieves high efficiency simultaneously in users search behaviours.

In this paper, we perform multi keyword search over encrypted data on the cloud to know the exact scope of such an application a survey was needed for user requirements and search behaviour. We found enabling the keyword search over the encrypted data is not an easy task some techniques allow the user to search over encrypted data securely through single keyword to retrieve documents of interest. This is insufficient as many users may tend to provide multiple keywords instead of one as their search interest. Recently methods have been proposed for multiple keyword search in cloud computing. In this method a binary index vector needs to be built for each document and each bit denotes whether the corresponding keyword is include in the document. The storing and updating index can be of substantial overhead especially when number of keywords is large. Thus the efficiency of secure multiple keyword search has large room for improvement for enhancing the system usability in cloud computing.

II. REVIEW OF RELATED LITERATURE

A. Privacy

The basic right to freedom is privacy it is an ability of an individual or group to seclude themselves or information about themselves and thereby express themselves selectively. The boundaries and the content of what is considered private and differ among cultures and individuals. In this paper we consider internet privacy to be one of the most important aspects of the internet. Internet privacy actually involves the right or mandate of personal privacy concerning the storing repurposing, provision to third parties and displaying the information pertaining to oneself via the internet. Privacy can entail either personally identifying information or non PII such as site visitors behaviour on the website, age and physical address and many more.

While dealing with the issue with internet privacy a person must first be concern with not only the technological implications such as damage property, corrupted files and the like but also with the potential of the implications of real life. One such implication which is rather commonly viewed as being one of the most daunting fears risks of the

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

internet, is the potential of identity theft although it is a typical belief that larger companies are the usual focus of identity thefts rather than individuals the recent report seem to show a trend opposing this belief specifically it was found that in 2007 “Internet Security Threat Reports” that roughly 93% of gateway attacks were targeted at unprepared home users. Gateway attack is an attack which is aimed not at stealing data immediately but actually getting access for future attacks.

B. Conceptual Framework

The basic concept of privacy preserving multi keyword search is to deliver a result similarity ranking to meet effective data retrieval need using coordinated matching of multi keywords we can constitute a multi cohesive environment for effectively capturing the relevance of outsourced documents to the query keywords and to use strategies to quantitatively evaluate the similarity measure through analysis investigating privacy and efficiency multiple experiments on the real world data show an introduction to low overhead on computation and communication

III. RESEARCH METHODOLOGY

Using multiple surveys with variety of users specific privacy requirements were generated.

1. Index Confidentiality and Query Confidentiality:
The under linked plain text information including keywords in index and query TF values of keywords stored in index and IDF values of keywords should be protected from cloud server.
2. Trapdoor Unlink ability:
The cloud server should not be able to determine whether to encrypted queries or trapdoors are generated from same search request.
3. Keyword Privacy:
The cloud server should not identify the specific keyword in query, index or document collection by analysing statistical information like term frequency. Searchable encryption scheme whose updating operation can be completed by cloud server only meanwhile reserving the ability to support multi keyword ranked search.

IV. SYSTEM ARCHITECTURE

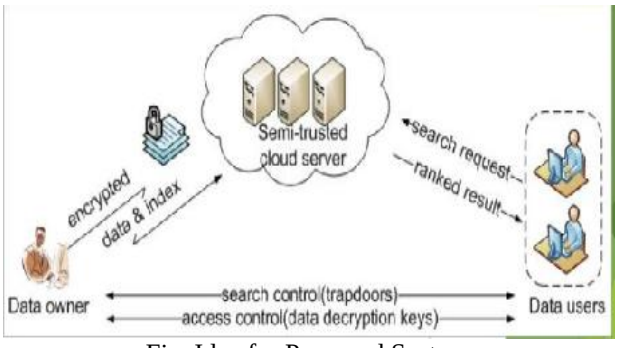

Fig. Idea for Proposed System

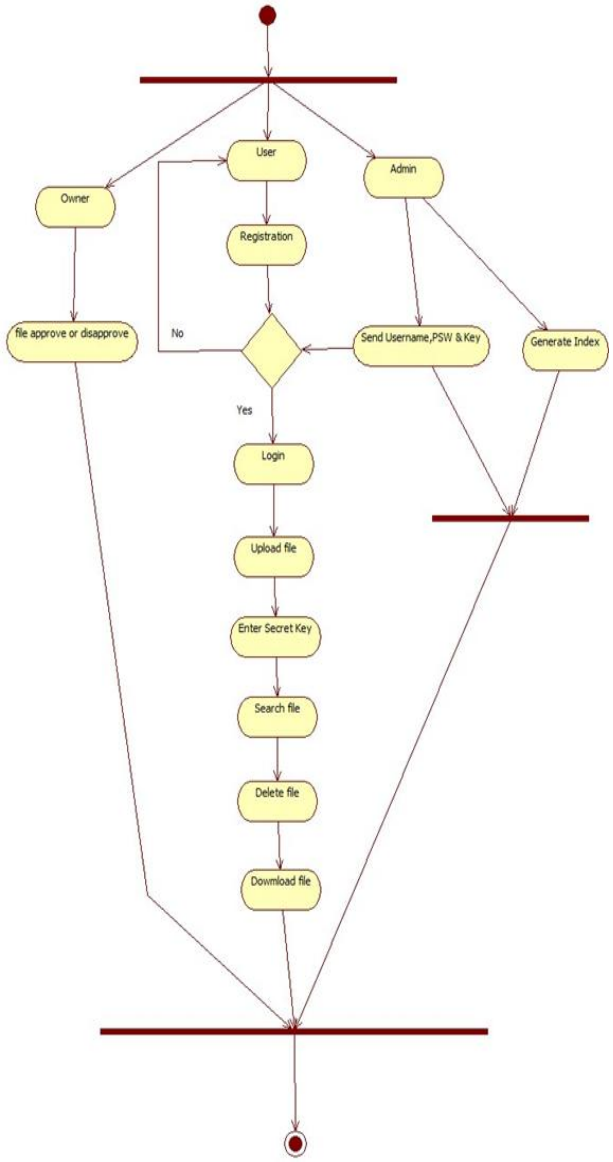

Fig. Basic Flow for the Project

V. DISCUSSIONS

The multiple keyword based structures tend to create dynamic search functionalities such as single keyword search, similarity search, multi keyword search, ranked search and many more. The propose system will have greater advantage at outsourcing sensitive information such as emails, personal health records and many more. Cloud service providers that keep the data for users may access users sensitive information with proper authorization.

The multiple schema based data will be primarily based on multi privacy requirements which will directly help the system in a dynamic manner. In depth analysis indicates multiple records are kept in the outsource system which may or may not be tracked in the current system. Multiple schemes such as BDMRS and COA are the most important schemes which needs to be handled.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 11, November 2016

CONCLUSION

This system creates the privacy preserving environment for better and formal access of data. The data privacy is important and sensitive data must always be encrypted. Thereby establishing a set of strict privacy requirements for a secure cloud data utilization system proposing an idea for an MRSE based on inner product computation and then give to significantly improved MRSE schemes to achieve stringent privacy requirements in two different threat models.

ACKNOWLEDGEMENT

We thank our colleagues who provided insight and expertise that greatly assisted the research. Then we thank our mentor Akash Bhojraj for guidance on the project. We also thank our **Prof. Nilesh Sable** for inspiring us to do this project and his comments on the manuscript.

REFERENCES

1) Chi Chen, Xiaojie Zhu, Peisong Shen, Jiankun Hu, Song Guo, Zahir Tari, and Albert Y. Zomaya, "An Efficient Privacy-Preserving Ranked Keyword Search Method" April 2016
2) S. Grzonkowski, P. M. Corcoran, and T. Coughlin, "Security analysis of authentication protocols for next-generation mobile and CE cloud services," in Proc. IEEE Int. Conf. Consumer Electron,2011, Berlin, Germany, 2011, pp. 83–87.
3) D. X. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Security Privacy, BERKELEY, CA, 2000, pp. 44–55.
4) D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT, Interlaken, SWITZERLAND, 2004, pp. 506–522.
5) http://www.slideshare.net/papithavelumani/privacy-preserving-multikeyword-ranked-search-over-encrypted-cloud-data-40291056?next_slideshow=1
6) http://www.slideshare.net/WINGZTECHNOLOGIESCHENNAI/privacy-preserving-multikeyword-ranked-search-over-encrypted-cloud-data-37727749
7) http://www.slideshare.net/CEGONTECHNOLGIES/multikeywordranked
8) ijcert.org/ems/ijcert_papers/V3I607.pdf