International Journal of Innovative Research in Engineering Science and Technology

APRIL 2018 ISSN 2320 –981X

AN EFFICIENT KEYWORD SEARCH OVER ENCRYPTED OUTSOURCED DATA

[1]Ms. Farheen.A.S, [2]Ms. Gayathri.T, [3]Ms. Lavanya.M

1.Department of Computer Science and Engineering,

Vivekanandha College of Engineering for
Women, Tiruchengode-637205
farheensamad1995@gmail.com

2.Department of Computer Science and Engineering,

Vivekanandha College of Engineering for
Women, Tiruchengode-637205
duraigayathri333@gmail.com

3.Department of Computer Science and Engineering,

Vivekanandha College of Engineering for
Women, Tiruchengode-637205
lavanyapackiya@gmail.com

CORRESPONDING AUTHOR:
Mr. R.T. Dinesh Kumar, M.Tech.,
Asst. Professor, Vivekanandha College of Engineering for Women,
Tiruchengode-637205
Email: dinesh@vcew.ac.in
Contact:9952762214

Abstract-In practice keyword-based search over encrypted data has become an important to in the current cloud computing scenario. In addition such keywords may have a certain grammatical relationship among them which reflect the importance of keywords from the user's perspective. However data encryption makes effective data utilization a challenging task. Traditional data utilization based keyword search on encrypted data is a difficult task. In this paper for the first time, we take the relation among query keywords into consideration and design a keyword weighting algorithm to show the importance of the distinction among them. By introducing the keyword weight to the search protocol design, the search results will be more in line with the user's demand. Keyword-based search schemas ignore the semantic representation information of user's retrieval, and cannot completely meet with user search intention. To better express the relevance between queries and files, we further introduce the TF-IDF rule when building trapdoors and the index. The datasets usually are encrypted before outsourcing to preserve the privacy. In particular, our scheme supports both dataset and keywords updates by using the sub-matrix technique. Experiments on the real-world dataset show that our proposed schemes are efficient effective and secure.

1. INTRODUCTION

Data owner encrypts and uploads data to cloud server, these data, which is stored as cipher text can be request and download by users. User who wants to server pose a query to cloud server, the query dealt in existing system is single keyword download data from cloud query, for which k-results are arrived, the results may not be much relevant to user and poses high communication cost. However, mobile cloud storage system faces challenges over the traditional encrypted search schemes. As mobile cloud devices have limited computing and battery capacities, there is a need for suitable and efficient encrypted search scheme is necessary for MCS. The mobile cloud storage requires high bandwidth and energy efficiency for data encrypted search scheme, due to the limited battery life and payable traffic fee. To overcome, these problems, we focus on the design of a mobile cloud scheme that is efficient in terms of both energy consumption and communication overhead, while keep meeting the data security requirements through wireless communication channels. To reduce high overhead, we propose an algorithm multi-keyword energy efficient search on encrypted cloud data.

2. EXISTING SYSTEM

Input central keywords with certain adjunct words as the query keywords when searching documents. The importance of each query keyword depends on the search intension of a user. So far many works have demonstrated the importance of keywords. The super-increasing sequence is to show the preference factors of keywords to indicate the importance of keywords in a query keyword set. However, users need to sort keywords according to their importance, which increases the users' input cost. Due to the lack of the super-increasing sequence, the last keyword the user inputs are more important than all the other keywords. The existing model built a user interest model for individual user by analyzing his search history. However, when inputting unusual keywords, it needs to re-build a new interest model. In this project, our use the grammatical relations as standards to show the weight of each keyword, and this enables users to retrieve relevant documents from the cloud based on their own interests.

2.1 Drawbacks of Existing System

Different kinds of access mechanism are not applied and so different client applications with varying processing capabilities need to execute the cloud data in same manner .Time limit is not discussed and so client like to access the data in same tariff for the whole period Correlated Authentication aspects with combination of both cloud storage provider application service provider and end user is not considered.

3. PROPOSED SYSTEM

In addition with all the existing system mechanism a correlated Authentication aspect with combination of the cloud storage provider application service provider and end user is also considered. In addition, time limit is provided to end user to access the Application Service Providers (ASPs). So at different time intervals, different kinds of tariffs can be applied to end users to access the service. Likewise, the security aspects provided by the cloud storage provider is also taken by ASPs to increase the security more. In addition, trusted third party authentication mechanism is included.

3.1 Advantages of proposed system

Different kinds of access mechanism are applied and so different client applications with varying processing capabilities need to execute the cloud data in same manner. Time limit is set and so client likes to access the data in different level for diverse time periods. Correlated Authentication aspects with mixture of both cloud storage provider application service provider and end user is also considered. Trusted third party authentication with no security violation is included.

4. PROJECT DESCRIPTION

To propose multi-keyword energy efficient search MKEE architecture for mobile cloud storage applications. MKEE achieves the efficiencies through employing and modifying the ranked keyword search as the encrypted search platform basis, which has been widely employed in cloud storage systems. Ranked keyword search procedure is modified to save the energy consumption of mobile devices, and proposed scheme simplifies the encrypted search procedure to reduce the traffic amount for retrieving data from encrypted cloud storage. MKEE is implemented with security enhancement based on popular TF-IDF.

5. SYSTEM ARCHITECTURE

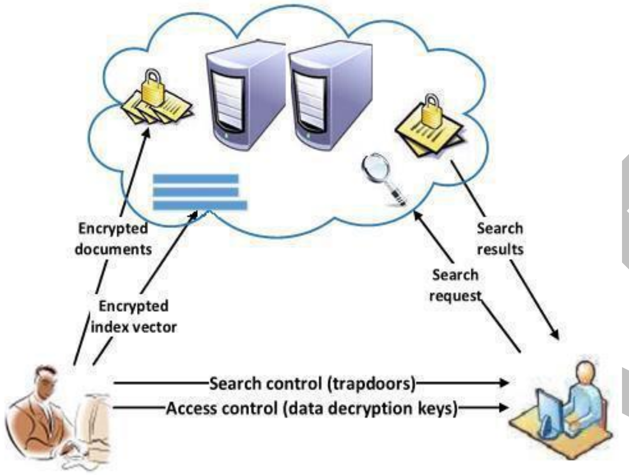

Fig.5.1 Cloud System Architecture

6. LITERATURE SURVEY

6.1 Problem Analysis A break in the clouds: Towards a cloud definition

This paper discusses the concept of Cloud Computing to achieve a complete definition of what a Cloud is, using the main characteristics typically associated with this paradigm in the literature. More than 20 definitions have been studied allowing for the extraction of a consensus definition as well as a minimum definition containing the essential characteristics. This paper pays much attention to the Grid paradigm, as it is often confused with Cloud technologies. We also describe the relationships and distinctions between the Grid and Cloud approaches.

6.2 Design of security solution to mobile cloud storage

The cloud storage owns advantages in pay for use and elastic scalability. However, the data security risk destroys the trust relation between the cloud service provider and user. A direct method to avoid this problem is to encrypt data before data stored in the cloud. Thus, without the decryption key, the leakage data cannot be decrypted. While the encryption technology is good, it is not always suitable for the mobile user. When using the mobile device, such as smart phone, to access the data that stored in cloud storage system, the performance issue should be considered, because the encryption scheme involves high workload. This paper is focus on the design of security solution to mobile cloud storage. It detailed the design principle, security function model, and typical deploy model. It also proposed a design case based on searchable encryption to guide the further research.

6.3 Impact of storage acquisition intervals on the cost- efficiency of the private vs. public storage

The volume of worldwide digital content has increased nine-fold within the last five years, and this immense growth is predicted to continue in foreseeable future reaching 8ZB already by 2015. Traditionally, in order to cope with the growing demand for storage capacity, organizations proactively built and managed their private storage facilities. Recently, with the proliferation of public cloud infrastructure offerings, many organizations, instead, welcomed the alternative of outsourcing their storage needs to the providers of public cloud storage services. The comparative cost-efficiency of these two alternatives depends on a number of factors, among which are e.g. the prices of the public and private storage, the charging and the storage acquisition intervals, and the predictability of the demand for storage. In this paper, we study how the cost-efficiency of the private vs. public storage depends on the acquisition interval at which the organization re-assesses its storage needs and acquires additional private storage. The analysis in the paper suggests that the shorter the acquisition interval, the more likely it is that the private storage solution is less expensive as compared with the public cloud infrastructure.

7.PROBLEM FORMULATION

7.1 System Model

Data owner upload data to cloud server. Data owner encrypts the data using popular encryption algorithm such as AES (Advance encryption standard) converts to cipher text and upload to cloud server. Consider a file set F= (F,1 F2, … Fn) containing the number of F files, the keywords index is also get a input from data owner. A table is created to store the file id and its corresponding keywords. TF table as our index and the cloud server calculates the relevance scores using the encrypted TF values. The ranking function of the relevance scores will be introduced in the subsection dedicated to the cloud server.

7.2 Threat Model

Data owner store their data on external servers that leads to increasing demands and

concerns for data confidentiality, authentication and access control. This concerns originate from the fact that cloud servers or usually operated by commercial providers which are very likely to be outside of the trusted domain of users. The system should prevent cloud server from learning the plaint text of the data files.

7.3 Design Goals

Extension of the Central Keyword. Our primary goal is to design a central keyword extension search scheme. When a user inputs some query keywords, our scheme can effectively and accurately locate and extend the semantic of the central keyword. The returned results should be relevant to both the multiple keywords that the user inputs and the extension keyword.

7.4 Notations and Preliminaries

Note that the cloud server is semi trusted, and the unwrap function can be processed by the server. Upon receiving the tuple Wrap(w) = (h1; h2), the server calls Unwrap to get user keywords, searches into the TF table, and then sends back the corresponding files. Cloud server calculates the relevance scores and returns top-k relevant files according to the searching query from data user.

$$Score(Ws, Fc) = \sum_{w \in Ws} \frac{1}{|F_c|} \times (1 + \ln f_{c,w})| \times \ln(1 + \frac{D}{f_w})$$

Ws is the keyword set to be searched; Fc is a certain file in the file set; fc, w denotes the TF of the keyword w in the file Fc; Fc is the total length of Fc; fw is the number of files containing the keyword w and D is the total number of files.

8. RELATED WORK

A number of different mechanisms have been proposed for security aspects in cloud computing. Ranked Searchable Symmetric Encryption (RSSE) technique, allow users to securely search over encrypted data through keywords. In applications such as outsource private databases, usually the (sole) owner of the outsourced data plays the role of the authority who generates such capabilities for users.

To support more complex queries, conjunctive keyword search schemes over encrypted data have been proposed. They can potentially support arbitrary query types CNF/DNF formulas, However, with exponential complexity. In this scheme complete file collection is scanned and a list of keywords is selected from building searchable index, which needs lots of computation for updating the searchable index when a user upload or deleting a file.

CONCLUSION

Multi-keyword energy efficient search scheme is implemented as an initial attempt to create a traffic and energy efficient encrypted keyword search tool over encrypted cloud data developed an efficient implementation to achieve an encrypted search in a cloud data. The security study of proposed system shows that it is secure enough for mobile cloud computing, while a series of experiments highlighted its efficiency. MKEE is slightly more time and energy consuming than keyword compared to traditional strategies featuring a similar security level.

REFERENCES

1.A. A. Moffat, T. C. Bell e,Feb-t al., Managing gigabytes: compressing and indexing documents and images. Morgan Kaufmann Pub, 1999. [8] D. Song, D. Wagner, and A. Perrig,"Practical techniques for searches on encrypted data," in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44– 55.

2. D. Boneh, G. Di Crescenzo, R.Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Advances in Cryptology-Eurocrypt 2004. Springer, 2004, pp. 506–522.

3.N.Nasurudeen Ahamed, "An Efficient method to group in the Social Media Using Micro Blogging Information",Internationl Journal of Engineering and Computer Science.Vol-6 ,issue -2,Feb -2017.

4.J.Oberheide, K. Veeraraghavan, E.Cooke, J.Flinn, and F.Jahanian, "Virtualized in-cloud security . services for mobile devices," in Proceedings of the First Workshop on Virtualization in Mobile Computing. ACM, 2008, pp. 31–35.

5.O. Mazhelis, G. Fazekas, and P. Tyrvainen, "Impact of storage acquisition intervals on the cost-efficiency of the private vs. public storage," in Cloud Computing(CLOUD), 2012 IEEE 5th International Conference on. IEEE, 2012, pp. 646–653.

6.D.Huang,"Mobilecloud computing,"IEEECOMSOC Multimedia Communications Technical Committee (MMTC) E-Letter, 2011.

7.X.Yu and Q.Wen, "Design of security solution to mobile cloud storage," in Knowledge Discovery and Data Mining. Springer, 2012, pp. 255–263.

8.L. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition,"ACMSIGCOMM Computer Communication Review, vol. 39, no. 1, pp. 50–55, 2008.

LIST OF FIGURES