International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 1, January - 2014

Secure and Efficient Search over Outsourced Data on Cloud

Rahul M. Harne
PG Student, Department of Computer Engg,
Smt. Kashibai Navale College of Engg
Pune, Maharashtra, India

S. P. Kosbatwar
Asst. Prof. Department of Computer Engg,
Smt. Kashibai Navale College of Engg
Pune, Maharashtra, India

Abstract

Cloud Computing is a service rather than product and it intends to make the internet the ultimate home of all computing resources. Cloud Computing uses data service outsourcing. Every individual organization wants to store data that has the limited storage capacity cloud computing outsource data to some Cloud Service Provider. For large data searching various techniques are keyword searchable encryption technique allow user to search data through keywords. In searchable encryption having problem searching go through every retrieving file and unnecessary network traffic. Paper defines and solves the problem of secure ranked keyword search over encrypted data on cloud. Paper proposes a Ranked keyword search system enhances system handling by search result and file retrieval correctness. In ranked keyword search system having data owner, data user and cloud server to effective data searching capabilities.

Index Terms: *Ranked search, searchable encryption, confidential data, cloud computing.*

1. Introduction

Cloud is the use of computing resources that is hardware and software that are delivered as a service over a network. The cloud makes it possible for people to access people information from anywhere at any time. One requirement is that people need to have an internet connection in order to access the cloud. This means that if people want to look at a specific document people have housed in the cloud, people must first establish an internet connection either through a wireless or wired internet or a mobile broadband connection. The benefit is that people can access that same document from wherever people are with any device that can access the internet. The devices could be a desktop, laptop, tablet or phone. Cloud computing provides many benefits in terms of low cost and accessibility of data.

The different types of clouds that are Public cloud, Private cloud, Community cloud, Hybrid cloud. As a home user or small organization owner, people will most likely use public cloud services.

Public Cloud - A public cloud can be accessed by any subscriber with an internet connection and access to the cloud space. Public cloud having applications, storage and other resources are available to the general public by a service provider of public cloud. The various services are free and also offered on a pay-per-use model. Generally, public cloud service providers like Amazon, Microsoft and Google own and operate and access only via Internet.

Private Cloud - A private cloud is established for a specific group or organization and limits access to just that group. Private cloud project requires degree of commitment to virtualized the business environment and it will require the organization to check decisions about existing resources. Private cloud is infrastructure operate on single organization and manage or hosted internally or externally.

Community Cloud - A community cloud is shared among two or more organizations that have similar cloud requirements, they managed internally or hosted internally or externally.

Hybrid Cloud - A hybrid cloud is essentially a combination of at least two clouds, where the clouds included are a mixture of public, private, or community. Hybrid clouds lack the security and certainty of in house applications.

Cloud computing involves delivery of services over internet. These services are divided in 3 terms – Platform as a Service (PaaS), Infrastructure as a Service (IaaS) and Software as a Service (SaaS).

PaaS - It is used for developing websites on the system without installing any software. It can be executed without any administrative expertise.

IaaS - If the revenue for cloud services primarily comes from charging for infrastructure. This System can be referred to as Infrastructure as a Service (IaaS). IaaS is the hardware and software that powers it all servers, storage, networks, operating systems.
SaaS - It is run by cloud service provider and mostly used by organizations. It is available to users through internet.
The cloud storage as a Service means that a third-party provider rents space on their storage to end users who lack the budget or capital budget to pay for it on their own. It is also ideal when technical personnel are not available or have inadequate knowledge to implement and maintain the storage infrastructure. Storage service providers are nothing new, but given the complexity of current backup, replication and disaster recovery needs, the services has become popular, among small and medium sized businesses. Biggest advantage of SaaS is cost savings. Storage is rented from the provider using a cost-per-gigabyte stored or cost-per-data transferred model. End user does not have to pay for infrastructure. People simple pay for how much they transfer and save on the provider server Cloud storage is becoming an increasingly attractive solution for organizations. That is because with cloud storage, data resides on the web, located across storage systems rather than at a designed corporate hosting site.
Cloud storage providers balance server loads and move data among various datacenters, ensuring that information is stored close and thereby available quickly. Storing data on the cloud is advantageous because it allows people to protect people data in case there is a disaster. People may have backup files of people critical information, but if there is fire or a hurricane wipes out people organization in this case having the backups stored locally does not help. Choosing a storage vendor can be a complex issue and how technology interacts with cloud can be complex. For instance, some product are agent based and the application automatically transfers information to the cloud via FTP but other employ a web front end and user has to select local files on their computer to transmit.

2. Issues in Cloud Computing

2.1. Security issues

Security issues are considered as most important issues in cloud computing. Following are some major security issues
Access Control – Unauthorized access may exist if security mechanism is not sufficient. As user data exist on cloud for long time so risk of criminal access is also more.
Authentification and Identification – Cloud serves having many clients that cloud allows one single instance of software to serve many clients. So there is problem of authentification and identification may happen.
Availability – User stores data on cloud and deletes its local copy. If service or data on cloud is not available due to some problem then it is hard to retrieve data.
Policy Control – The Cloud is mixed, which means that different Cloud servers may have different mechanisms to ensure the client's data security.
Audit – Cloud service provider has control of data. There is possibility that cloud service provider can alter data without client's permission. So, there is need of audit. [3]

2.2. Privacy issues

Privacy issues include protection of identity information, transaction histories and sensitive data. Idea of cloud computing is store user data on shared infrastructure. So, there is risk of unauthorized access. Following are some privacy issues
Unauthorized secondary usage – Users data can be utilized by cloud service provider. Unauthorized usage of data may cause serious security problems, which becomes one significant concern.
Lack of user control – As user no longer possesses its data. Data is not transparent to users which means user have lost control over data. So there is need of protection mechanism.
Unclear responsibility – There is one problem related to the privacy that it is sometimes unclear about which CSP is responsible for privacy protection, detecting who use and modify the user data or ensuring user data privacy requirements. [3]
Solution for ensuring data safety is retrieving the data from cloud owner, search cloud data using ranked search enhances system handling by search result and file retrieval correctness.

3. System Model

Cloud data storage consists of 3 entities which are used in the searching of outsourced data on cloud server with Ranked keyword search.
Data User – Who has large amount of data which is to be stored on cloud and send request to cloud server.
Cloud Server – This has computing resources to manage cloud data.
Data Owner – Challenge cloud server to check the correctness of data on behalf of user and indexing outsource data.

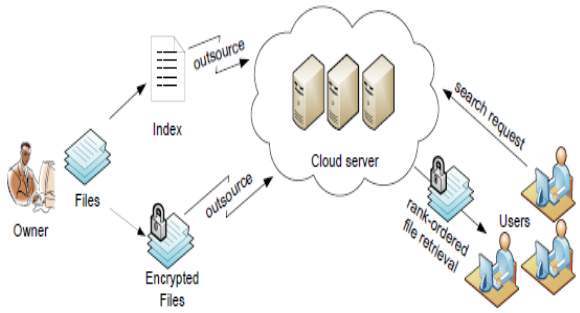

Fig 1. Architecture for search over cloud data.

Ranked keyword search on the cloud server and data owner with user request send. Ranked based symmetric encryption scheme is based on ranked keyword search. The user send request to cloud server to search file from data owner. In data owner files are in the indexing order to outsource data use score calculation to find hash index for each file stored in the cloud server. The outsource data from data owner and use encryption and use one to many order preserving encryption scheme to retrieving the file. [1]

4. Related Work

In data searching different techniques were used to provide searching to cloud data but there are some disadvantages of this system. Common methods for protecting user data files include encryption prior to storage and provide indexing in ranked searching, user authentication procedures prior to storage or retrieval, and building secure channels for data transmission.

C. Wang and W. Lou are motivate and solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored encrypted data in Cloud Computing. They first give a basic scheme and show that by following the same existing searchable encryption framework, it is very inefficient to achieve ranked search. Then appropriately weaken the security guarantee, choice to the newly developed crypto primitive OPSE, and derive an efficient one-to-many order-preserving mapping function, which allowing the effective RSSE to be designed. Through security analysis, they show that proposed solution is secure and privacy-preserving, while correctly realizing the goal of ranked keyword search. Extensive experimental results demonstrate the efficiency of our solution. [1]

D. Song, D. Wagner, and A. Perrig are providing provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the cipher text. They provide query isolation for searching that the untrusted server cannot learn anything more about the plaintext than the search result. They provide good control searching, so that untrusted server could not search for an arbitrary word without the user's authorization. They also support secrete queries, so that the user may ask the untrusted server to search for a secret word without revealing the word to the server. The algorithms presented are simple, fast (for a document of length n, the encryption and search algorithms only need O(n) stream cipher and block cipher operations), and introducing no space and communication surface, and hence these use practically today. [6]

E.-J. Goh formally define a secure index and formulate a security model for indexes known as semantic security against adaptive chosen keyword attack (ind-cka). They also develop an efficient ind-cka secure index construction called z-idx using pseudo-random functions and Bloom filters and show how to use z-idx to implement searches on encrypted data. This search scheme is most capable encrypted data search scheme, it provides O(1) search time per document and handles compressed data, variable length words and boolean and certain regular expression queries. The techniques developed can be used to build encrypted searchable audit logs, private database query scheme and accumulated hashing scheme and secure set membership test. [7]

5. Proposed System

In the proposed system, we will make use of multi cloud storage to upload clinical data on the cloud. If one of the server gets corrupted then file can be retrieve from other servers, the backup server is used for the retrieval of the file. As Cloud Computing becomes prevalent, more and more sensitive information are being centralized into the cloud, such as emails, personal health records, company finance data, and government documents, etc.

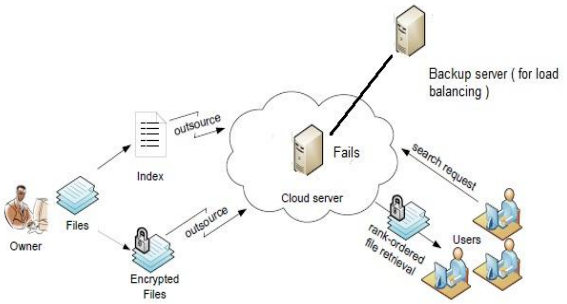

Fig 2. Architecture of search over cloud data with load balancing system

The fact that data owners and cloud server are no longer in the same trusted domain may put the outsourced unencrypted data at risk: the cloud server may leak data information to unauthorized entities. It follows that sensitive data has to be encrypted prior to outsourcing for data privacy and combating unsolicited accesses. If the server fails, there is no backup available then no one can handle this work, this is main problem in this system and it is not efficient because there is no facility of load balancing.

6. Conclusion

The existing searchable encryption framework, it is very inefficient to achieve ranked search data and appropriately weaken the security guarantee. The proposed system uses ranked keyword search technique for searching data on cloud server and also uses load balancing technique for the security. The more efficient searchable technique and backup service and security.

7. References

[1] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data" IEEE Transaction On Parallel and Distributed System, VOL. 23, NO. 8 Aug, 2012.
[2] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," Proc. IEEE INFOCOM '11, 2011.
[3] Ziyuan Wang, "Security and privacy issues within the Cloud computing", 2011 International Conference on Computational and Information Sciences.
[4] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy Keyword Search over Encrypted Data in Cloud Computing," Proc. IEEE INFOCOM '10, 2010.
[5] A. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order-Preserving Symmetric Encryption," Proc. Int'l Conf. Advances in Cryptology (Eurocrypt '09), 2009.
[6] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy'00, 2000.
[7] E.-J. Goh, "Secure indexes," Cryptology ePrint Archive, Report 2003/216, 2003, http://eprint.iacr.org/.
[8] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000
[9] Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Int'l Conf. Applied Cryptography and Network Security (ACNS '05), 2005.
[10] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions," Proc. ACM Conf. Computer and Comm. Security (CCS '06), 2006