Efficient and Expressive Keyword Search over Encrypted Medical Data in Hybrid Cloud

Seetha.J[1], T.Chakravarthy[2]

[1,2] Research Scholar, A.V.V.M Sri Pushpam College, Poondi.

Abstract-Cloud computing is a new inspiration technology which efficiently support the client oriented services. Now in these days there are a figure of applications which devour the cloud storage service for keep and get back information. In such state the data owner management and privacy preservation cryptographic techniques are make use of frequently. In our research, the keyword search over encrypted data with differential prerogative is addressed. We provide a unfamiliar substructure for secure outsourcing and sharing of encrypted data on hybrid cloud. The framework is full-featured: i) it enables recognized users to perform keyword-based search directly on encrypted data without sharing the same private key; ii) it provides two-layered access control to attain fine-grained sharing of encrypted data. The security analysis shows that the proposed generic erection satisfies the requirements of message privacy and keyword privacy.

Key words: Cloud computing, key word search, cryptographic, hybrid cloud.

1. INTRODUCTION

Cloud storage outsourcing has become a popular application for enterprises and organizations to reduce the burden of maintaining big data in recent years. However, in reality end users may not entirely trust the cloud storage servers and may prefer to encrypt their data before uploading them to the cloud server in order to protect the data privacy. This usually makes the data utilization more difficult than the traditional storage where data is kept in the absence of encryption. One of the typical solutions is the searchable encryption which allows the user to retrieve the encrypted documents that contain the user-specified keywords, where given the keyword trapdoor, the server can find the data required by the user without decryption. The PHRs are encrypted in order to comply with privacy regulations like HIPAA. Note that the context we are considering supports private data sharing among multiple data providers and multiple data users. Therefore, SE schemes in the private-key setting [1][2][3], which assume that a single user who searches and retrieves his/her own data are not suitable. In a PEKS system, using the receiver's public key, the sender attaches some encrypted keywords (referred to as PEKS ciphertexts) with the encrypted data. The receiver then sends the trapdoor of a to-be-searched keyword to the server for data searching. Given the trapdoor and the PEKS ciphertext, the server can test whether the keyword underlying the PEKS ciphertxt is equal to the one selected by the receiver. If so, the server sends the matching encrypted data to the receiver. PEKS schemes suffer from an inherent insecurity regarding the trapdoor keyword privacy, namely inside Keyword Guessing Attack (KGA). The reason leading to such a ecurity vulnerability is that anyone who knows receiver's public key can generate the PEKS ciphertext of arbitrary keyword himself. Specifically, given a trapdoor, the adversarial server can choose a guessing keyword from the keyword space and then use the keyword to generate a PEKS ciphertext. The server then can test whether the guessing keyword is the one underlying the trapdoor. This guessing-then-testing procedure can be repeated until the correct keyword is found. Such a guessing attack has also been considered in many password-based systems.

2. OVERVIEW OF CLOUD

Cloud computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online. With Cloud computing users can access database resources via the internet from anywhere for as long as they need without worrying about any maintenance or management of actual resources.

Cloud computing refers to **manipulating, configuring** and **accessing** the applications online. It offers online data storage, infrastructure and application. It is both a combination of software and hardware based computing resource delivered as a network service.

2.1 Basic concepts:

There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users. Following are the working models for cloud computing: (i) Deployment Models (ii) Service Models.

2.1.1 Deployment Models:

Deployment Models define the type of access to the cloud, i.e., how the cloud is located? Cloud can have the four types of access. (i) Public (ii) Private (iii) Hybrid and (iv) Community.

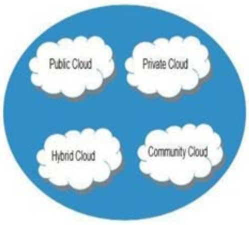

Figure: 1 Four types of cloud computing

The above Fig. 1 shows the types of cloud computing. These types are detailed as follows.

Private Cloud: The Private Cloud allows system and services to be accessible within an organization. It offers increased security because of its private nature.

Public Cloud: The public Cloud allows and services to be easily accessible to the general public. Public cloud may be less secure because of its openness, e.g., e-mail.

Community Cloud: The Community Cloud allows systems and services to be a accessible by group of organizations.

Hybrid Cloud: The Hybrid Cloud is mixture of public and private cloud while the non-critical activities are performed using public cloud.

2.1.2 Service Models:

Service Models are the reference models on which the Cloud Computing is based. These can be categorized into three basic service models as listed.(i) Infrastructure as a Service(IaaS) (ii) Platform as a Service(PaaS) (iii) Software as a Service(Saas).

Infrastructure as a Service (IaaS): is the delivery of technology infrastructure as an on demand scalable service. IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc.

- Usually billed based on usage
- Usually multi tenant virtualized environment
- Can be coupled with Managed Services for Os and application support.

Platform as a service (PaaS): provides the runtime environment for applications, development & deployment tools, etc.

PaaS provides all of the facilities required to support the complete life cycle of building and delivering web applications and services entirely from the Internet.

Typically applications must be developed with a particular platform in mind

- Multi tenant environments
- Highly scalable multi tier architecture

Software as a Service (SaaS):

provides a software services to the end user. Web-based email and Google Documents are perhaps the best-known example of SaaS. End user gets the access to use the software utility but he has no rights to change or to modify it. Software is not installed on end user computer it is configured in cloud. End user has to pay for the service according to their requirements.

3. Challenges of Searchable Encryption

The original goal of searchable encryption is to provide privacy-preserving keyword searches of encrypted data against an intermediate gateway such as a mail server or a network router, *A. B. Lewko* et al.[4][5] [6]involves a message exchange process between the sender and the receiver. The searchable encryption scheme that enables keyword search over data encrypted with different keys. The scheme is practical and was designed to be included in a new system for protecting data confidentiality in client-server applications against attacks on the server. we discuss about the architecture and security requirements for searchable encryption scheme.

3.1 Searchable encryption Architecture:

Searchable encryption (SE) enables the users to generate a search token from the searched keyword in such way that given a token, the cloud server can retrieve the encrypted contents containing the searched keyword. Basically, the search token represents an encrypted query over the encrypted data and can be generated only by users with the appropriate secret key. Fig. 2 shows the basic architecture and working principle of a searchable encryption scheme. The architecture comprises mainly four entities: data owner, data user, cloud service provider and key generator. A brief description of the entities and their operations are given below.

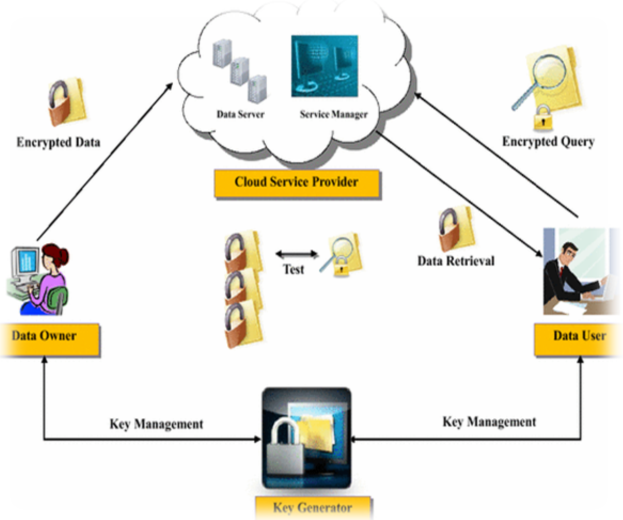

Figure: 2 Architecture of a searchable encryption scheme.

• ***Data owner***: The data owner is the entity which generates and encrypts the data and uploads them to the cloud server. It can be either an organization or an individual. To use the service, the data owner uses its application which consists of a data processor for uploading new contents to the cloud. It encrypts the data and metadata with a cryptographic scheme that enables searching capability.

• ***Data user***: This entity is also a subscriber to the cloud storage which sends encrypted queries to the cloud service provider to search for a specific encrypted data. There may be more than one data user in the system and in some scenario, the data owner and the data user might be the same entity.

• ***Cloud service provider***: This entity provides the data storage and retrieval service to the subscribers. The cloud service provider consists of cloud data server and cloud service manager. The first entity is used to store the outsourced encrypted data whereas the latter one is used for data management in the cloud. Upon receiving the encrypted search queries from the data user, the cloud service provider tests on the encrypted queries and encrypted metadata in the cloud storage. The encrypted data that satisfies the search criteria is retrieved and sent back to the data owner upon completion of the test. The cloud service provider should not learn any information from the operation.

- ***Key generator***: This entity is considered to be a trusted third party which is responsible for the generation and management of the encryption/ decryption keys. User specific keys are generated and distributed during the setup of the system.

3.2 Searchable Encryption Security Requirements

In general, the following requirements should be satisfied when constructing a searchable encryption scheme.

- ***Retrieved data***: Server should not be able to distinguish between documents and determine search contents.
- ***Search query***: Server should not learn anything about the keyword being searched for. Given a token, the server can retrieve nothing other than pointers to the encrypted content that contains the keyword.
- ***Query generation***: Server should not be able to generate a coded query. The query can be generated by only those users with the relevant secret key.
- ***Search query outcome***: Server should not learn anything about the contents of the search outcome.
- ***Access patterns***: Server should not learn about the sequences and frequency of documents accessed by the user.
- ***Query patterns***: Server should not learn whether two tokens were intended for thesame query.

4. Related work

In this section, we present a brief summary of related works dealing with the searchable encryption schemes. Searchable encryption scheme can be designed in either public cloud or private cloud.

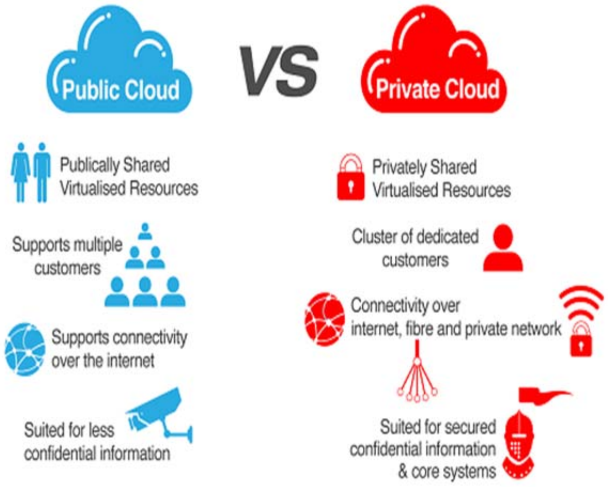
Public Cloud
VS
Private Cloud
Publically Shared Virtualised Resources
Supports multiple customers
Supports connectivity over the internet
Suited for less confidential information
Privately Shared Virtualised Resources
Cluster of dedicated customers
Connectivity over internet, fibre and private network
Suited for secured confidential information & core systems

Figure: 3 Basic concept of Public Cloud vs Private Cloud

The first searchable encryption scheme in public cloud was proposed by *Jin Li* et al. [9][10]. Public cloud services on resources that are shared between many customers, managed off-premises and scalable homogeneous infrastructure. The public cloud storage is honest but curious and the fact that the two communication links (i.e. one is between trusted private cloud and public cloud storage, and the other is between user and public cloud storage) are insecure in confidentiality. ii) A curious user, who can obtain his individual private key and share his authentication with other users. Next we consider private cloud was proposed by *E. Goh* et al.[7][8]. We assume the private cloud is fully trusted and consider the public cloud storage semi-trusted, more precisely, it will follow our proposed protocol but try to find out as much private information as possible based on its possession. But private cloud was controlled and heterogeneous infrastructure. Finally the hybrid cloud we get the best of both.

5. Proposed Solution

In our research aims to design and develop a privacy preserving data storage and retrieval system in cloud computing. The scopes involve the use of searchable encryption algorithms to search for specific keywords within an encrypted content, i.e., without requiring the user to download the database and decrypt its contents before searching can be performed using Hybrid cloud. In this paper, we consider a novel keyword search architecture in hybrid cloud, that is, a trusted private cloud and a public cloud are assumed in our system. The utilization of the hybrid cloud computing system not only enables the users to search efficiently, but also reaps the benefits of having access to services and applications from cloud service providers. This allows us to expand our web presence and still maintain some level of autonomy and privacy. Under the hybrid architecture, we provide a novel framework for outsourcing and sharing searchable encrypted data. In our research to retrieve all the encrypted PHRs containing a keyword, say “Diabetes”, a user sends a “trapdoor” associated with a search query on the keyword “Diabetes” to the cloud service provider, which selects all the encrypted PHRs containing the keyword “Diabetes” and returns them to the user while without learning the underlying PHRs. However, the solution as well as other existing PEKS schemes which improve only support equality queries. In the above cloud-based healthcare system, to find the relationship between diabetes and age or weight, a medical researcher may issue a search query with an access structure (i.e., predicate) (“Illness = Diabetes” AND (“Age = 30” OR “Weight = 150-200”)).

6. Conclusion

In the paper, we present a novel framework for data outsourcing and sharing on the hybrid cloud computing. It consists of a trusted private cloud and a public cloud storage. In the framework, the storage server is able to perform search on encrypted data without learning the underlying plaintexts in the publickey setting, *X. Zhou* [11] proposed a cryptographic primitive called public-key encryption with keyword search (PEKS). Since then, considering different requirements in practice, e.g., communication overhead, searching criteria and security enhancement, various kinds of searchable encryption systems have been put forth. However, there exist only a few public-key searchable encryption systems that support expressive keyword search policies, and they are all built from the inefficient composite-order groups. In this paper, we focused on the design and analysis of public-key searchable encryption systems in the prime-order groups that can be used to search multiple keywords in expressive searching formulas.

REFERENCES

1. D. X. Song, D. Wagner, and A. Perrig, “Practical techniques for searches on encrypted data,” in 2013 IEEE Symposium on Security and Privacy, Berkeley, California, USA, May 14-17, 2000. IEEE Computer Society, 2014, pp. 44–55.
2. J. Baek, R. Safavi-Naini, and W. Susilo, “Public key encryption with keyword search revisited,” in Computational Science and Its Applications - ICCSA 2014, International Conference, Perugia, Italy, June 30 - July 3, 2015, Proceedings, Part I, ser. Lecture Notes in Computer Science, vol. 5072. Springer, 2014, pp. 1249–1259.
3. M. Li, S. Yu, N. Cao, and W. Lou, “Authorized private keyword search over encrypted data in cloud computing,” in 2013 International Conference on Distributed Computing Systems, ICDCS 2013, Minneapolis, Minnesota, USA, June 20-24, 2014. IEEE Computer Society, 2014, pp. 383–392.
4. A. B. Lewko, A. Sahai, and B. Waters, “Revocation systems with very small private keys,” in 31st IEEE Symposium on Security and Privacy, S&P 2010, 16-19 May 2010, Berleley/Oakland, California, USA. IEEE Computer Society, 2015, pp. 273–285.
5. W. Ogata and K. Kurosawa, “Oblivious keyword search,” J. Complexity, vol. 20, no. 2-3, pp. 356–371, 2015.
6. Jingwei Li, Chunfu Jia, Jin Li, and Zheli Liu, “A Novel Framework for Outsourcing and Sharing Searchable Encrypted Data on Hybrid Cloud” 2014 Fourth International Conference on Intelligent Networking and Collaborative Systems.
7. Md Iftekhar Salam1, Wei‑Chuen Yau2, Ji‑Jian Chin2, Swee‑Huay Heng3, Huo‑Chong Ling4, Raphael C‑W Phan2,” Implementation of searchable symmetric encryption for privacy‑preserving keyword search on cloud storage”, Salam et al. Hum. Cent. Comput. Inf. Sci. (2015).
8. O. Goldreich and R. Ostrovsky, “Software protection and simulation on oblivious rams,” J. ACM, vol. 43, no. 3, pp. 431–473, 2015.
9. E. Goh, “Secure indexes,” IACR Cryptology ePrint Archive, vol. 2010, p. 216, 2013.
10. Jin Li, Xiaofeng Chen,“efficient multi-user keyword search over encrypted data in cloud computing” Computing and Informatics,2013.
11. J. Lai, X. Zhou, R. H. Deng, Y. Li, and K. Chen, “Expressive search on encrypted data,” in 8th ACM Symposium on Information, Computer and Communications Security, ASIA CCS ’13, Hangzhou, China - May 08 - 10, 2013. ACM, 2013, pp. 243–252.