SECURE RETRIEVAL OF FILES USING HOMOMORPHIC ENCRYPTION FOR CLOUD COMPUTING

R. Sharmila[1]

[1]*PG Scholar, Computer Science and Engineering, R.M.K. Engineering College, Tamil Nadu, India*

Abstract

Clouds allow users to store data and access can be made anywhere, any time by using any device. Highly sensitive information such as business documents, medical records and personal information may be stored in a cloud. Security and privacy are thus very important issues in cloud computing. To keep user data confidential from an untrusted Cloud Service Provider and third parties, a natural way is encryption. The data decryption key should be disclosed only to users who have been authorized. Users can search their files using keywords in the cloud. In existing literature many schemes have been proposed. In this paper, a new technique is described: Multi-keyword searching using homomorphic encryption. It is an algorithm which performs operations on encrypted data which will provide results without decrypting that data. It provides privacy for user querying patterns and user data. It allows Cloud Service Providers to perform operations on the encrypted data. The Cloud Service Provider is unaware of the files and keywords stored in the cloud. Ranking is used for efficient and fast retrieval of the desired files. Ranks will be assigned to files based on the frequency of access of the files.

Keywords: *Homomorphic encryption, Multi-keyword searching, Cloud Service Provider, Ranking.*

----------***----------

1. INTRODUCTION

Cloud computing is an important trend in the present scenario. However, there is no standard definition for cloud computing, nevertheless, the National Institute for Standards and Technology (NIST) provides the following workable definition: "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

NIST mentions the essential characteristics of cloud computing as on demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. The service models commonly employed in cloud computing are: Software as a service (SaaS), Platform as a service (PaaS), and Infrastructure as a service (IaaS). The deployment models normally used in cloud computing include: Private cloud, Public cloud, Community cloud, Hybrid cloud [1].

One of the most fundamental services provided by the cloud is Storage as a Service. In this service, data security, data privacy and data retrieval are considered as the main issues. Even though the cloud has challenging issues such as security, privacy, attacks, data loss and leakage, it has attractive benefits: capability to share resources and maximize their utilization, provision for pay as you use, reduced infrastructure and maintenance costs, and accessibility and flexibility.

Cloud computing is capable of providing flexible infrastructure. Hence, users need not own the infrastructure. Cloud computing supports high scalability and multi-tenancy. The payment for using the services may be made on a pay-per-use basis. Cloud users should be aware of trusted and untrusted Cloud Service Providers before storing their data in cloud. It is important for the data owner how the Cloud Service Provider will regulate access to data and keep it secure. Cloud Service Providers must make sure that the customer's personal information is well secured from third parties.

Assume that a user wants to store his files in a cloud and later he wants to retrieve specific files. The best approach is keyword-based searching for retrieving files. If the user stores his data without any protection mechanism, there is a chance for attacks such as internal and external attacks. User data and querying patterns should not be revealed to any untrusted Cloud Service Providers or unauthorized attackers. It must be kept secured. Therefore encrypted files are stored in the cloud. Along with the files, keywords are also encrypted and stored. If simple encryption is used then the Cloud Service Provider will return all the files related to the keyword because the encryption is not searchable. The encryption used should be searchable and therefore users can retrieve their desired files.

Different schemes were introduced for keyword searching over encrypted data. It is efficient and supports searching on the encrypted data. To search and retrieve securely Searchable Encryption was developed. It allows the Cloud Service Provider to verify whether the keyword specified by a user is contained in a file. The Cloud Service Provider is unaware of the keywords and files.

Many techniques support just single keyword search. Users have to provide exact keywords. Due to single keyword

search, retrieval of particular files was limited. Multi-keyword searching techniques were also introduced; however, they do not provide the desired security levels for user data stored in a cloud and user querying patterns when retrieving files.

To overcome the above drawbacks, we use Multi-Keyword searching using Homomorphic Encryption. It is a novel technique which aims to provide data confidentiality in the cloud. Homomorphic encryption is an algorithm which performs operations on the encrypted data, User data and query patterns can be maintained securely.

The rest of the paper is organized as follows: Related works are reviewed in Section 2. The proposed system is described in Section 3. Performance is explored in Section 4. Finally, conclusions are given in Section 5.

2. RELATED WORKS

Q.Liu et al. [5] proposed the Secure and Privacy Preserving Keyword Search (SPKS) scheme which enables the Cloud Service Provider not only to determine which files contain certain keywords specified by the user, but also to participate in the partial decipherment to get the result of the decipherment before returning the search results. Thus it reduces both the communication and computational overhead in decryption for the user, on the condition of preserving user data privacy and user querying privacy. They had implemented this scheme with single keyword searching.

To retrieve a desired file many schemes have been developed in recent years. C.Orencik and E.Sava [6] described the Private Information Retrieval (PIR) protocol. Privacy-preserving ranked keyword search scheme based on Private Information Retrieval (PIR) is used which allows multi-keyword queries with ranking capability and provides security.

After searchable encryption, fuzzy-keyword search was also introduced. Fuzzy-keyword search is a spell check mechanism. J.Li et al. [3] discussed Wildcard-based fuzzy-keyword set constructions with edit distance. Fuzzy-keyword search greatly enhances system usability by returning the matching files when users searching inputs exactly match the pre-defined keywords or the closest possible matching files based on keyword similarity semantics, when exact match fails. Wildcard-based fuzzy keyword set construction with edit distance is used by the authors and it does not support sequences of keywords.

W.Zhou et al [8] defined K-grams based fuzzy-keyword sets. In their article, two separate servers which cannot communicate with each other are used to provide security. There are search servers and storage servers. In this scheme, building query requests includes three procedures: generating fuzzy keyword set according to the building coefficient, generating trap doors, and generating an index. It should be noted that the size of the fuzzy keyword set should be controlled.

3. PROPOSED WORK

3.1 System Model

Figure-1 illustrates the architecture of cloud storage. It consists of the following entities:

Data Owner: Data owner will encrypt his files and keywords and store it in the cloud. When he is interested in retrieving a file using multi-keyword searching he can retrieve it. He can store his files in public clouds as well as in private clouds.

Other Users: They are the users to whom the data owner has given rights to access the files stored in a public cloud. The authentication details of these users are verified using cloud servers.

Key Server: It is a reliable server which stores the keys for the encrypted files in the cloud. Files and keywords are stored along with signature of the files. It provides the decryption key for a file after verifying the signature of the file name.

Cloud Server: It is a server where the encrypted files and the keywords are stored. It verifies the authentication of the users. It allows the Cloud Service Provider to perform operations on the encrypted data.

Cloud Service Provider: The data owner has to choose reliable Cloud Service Providers who will keep the clients information secure and must prevent it from any attacks. When the data owner specifies the keywords, the Cloud Service Provider will verify the keywords and provide it to the data owner. Both the Cloud Service Provider and the cloud server are unaware of the files and the keywords.

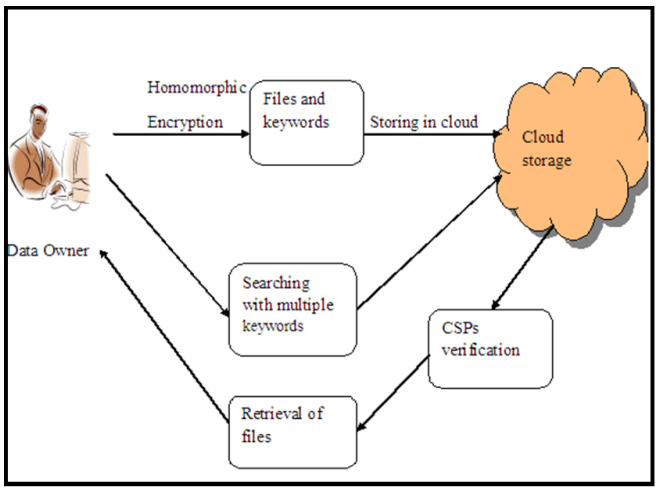

Fig-1. Architecture of Proposed Work

3.2 Working Methodology

Homomorphic Encryption

Definition: It allows specific types of computations to be carried out on the cipher text and to obtain an encrypted

result which when decrypted matches the result of operations performed on the plaintext.

Before it is sent, the data in the cloud is encrypted with homomorphic encryption and operations are executed in the encrypted data and the results are decrypted, it is same as the operations performed on the original data. Homomorphic encryption cryptosystem provides security and data confidentiality.

In homomorphic encryption, there is just one operation on the plaintext that has a corresponding operation on the cipher text. For example, plain RSA has that property. Suppose cipher text c1 is the encryption under a public key pk of plaintext m1, and c2 is the encryption under the same key of m2; that is

$$c1 = enc(pk, m1) \text{ and } c2 = enc(pk, m2).$$

Then multiplying the cipher texts results in something which, when decrypted, is identical to the result of multiplying the two plaintexts. That is, if sk is the decryption key corresponding to pk, then

$$m1 \times m2 = dec(sk, c1 \times c2) [9].$$

An algorithm is homomorphic if it is made up of addition, subtraction, multiplication functions. In the proposed system, we use multiplication while encrypting the data. This algorithm consists of the following steps:
1. Key generation
2. Encryption
3. Decryption

The data owner encrypts his files and keywords using this algorithm and stores it in cloud storage. The number of times the files and keywords has to be encrypted may be specified. When the data owner requires his file to be retrieved, he can request the Cloud Service Provider. The Cloud Service Provider will perform computations on the encrypted data without knowing anything of the files and the keywords. It will send back the results to the data owner. The data owner can decrypt it.

Paillier cryptosystem supports the both additive and multiplicative homomorphic encryption. In proposed system we use multiplication property for encryption. The algorithm steps are defined in [10].

The owner can store his files in the private cloud as well as in the public cloud. If he wishes other users to access files which are publicly stored he can give rights to them. The owner will send the secret key for authorized users. Their access rights are verified by the cloud sever. The owner should be aware of the revoked users.

Multi-keyword searching is used to retrieve a desired file from a database. When a user wants to retrieve a file, he can enter multiple keywords. The Cloud Service Provider could determine which all files that contain the specified keyword without revealing anything about the contents of the document or the keyword searched. The Cloud Service Provider will verify the keywords specified by the user and the keywords already stored in cloud.

The owner will be asked for the secret key and the number of times the encryption used during the encryption. By defining those,he will be able to retrieve his file. The time consumption for retrieving the desired file should be less. Therefore ranking scheme is used. While searching with keywords, it is difficult to decide which documents are the most relevant. We can build an index for frequently occurring keywords. Ranks are assigned to the files based upon a frequent access to the file. The file that contains the highest ranks and the keywords specified by the data owner can be retrieved quickly. The top-most files are retrieved and the data owner can download it.

4. PERFORMANCE EVALUATION

In this section, the performance of our technique is evaluated. The experiment is implemented using java language and tomcat server deployment. The time consumption for retrieving a file is less using ranking scheme. **Chart-1** defines the performance of retrieval of file using Homomorphic encryption.

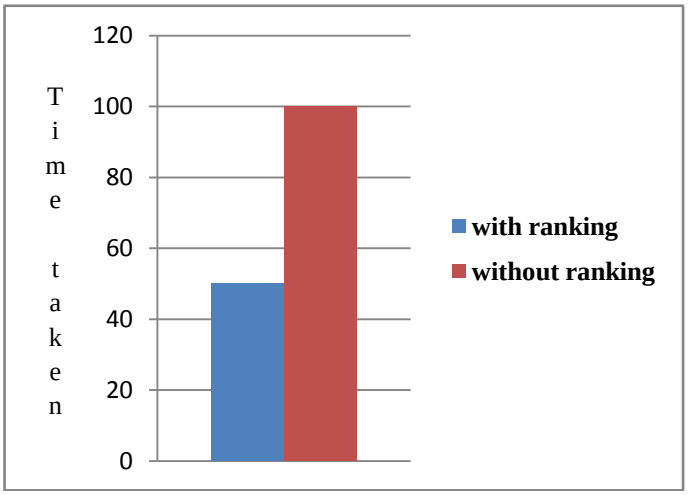

Chart-1 Retrieval of files

5. CONCLUSIONS

This paper describes the problem of retrieving files from cloud storage and keeping the files secure from untrusted Cloud Service Providers and third parties. The homomorphic encryption algorithm aims to give privacy for user data and querying patterns. Users can encrypt their files along with keywords using this algorithm. It provides confidentiality for user querying patterns and privacy for user data. The Cloud Service Providers are unaware of the files and the keywords. Ranking scheme improves the retrieval of files. It reduces the time taken to retrieve the files. Finally, security and confidentiality of data is maintained in the cloud. Thus, homomorphic encryption has many advantages.

Multi-keyword searching is used to retrieve a desired file from a database. When a user wants to retrieve a file, he can enter multiple keywords. The Cloud Server could determine which all files contain the specified keyword without revealing anything about the contents of document

REFERENCES

[1]. The NIST Definition of Cloud Computing, NIST Special Publication 800-145, 2011.
[2]. B. Hayes, American Scientist- The Magazine of Sigma Xi, The Scientific Research Society, 2012.
[3]. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, Fuzzy Keyword Search over Encrypted Data in Cloud Computing, Proceedings of IEEE, INFOCOM, pages:1– 5, 2010.
[4]. M. Li, S. Yu, N. Cao and W. Lou, Authorized Private Keyword Search over Encrypted Data in Cloud Computing, Proceedings of Distributed computing systems (ICDCS), pages:383 – 392, 2011.
[5]. Q. Liu, G. Wang, and J. Wub, Secure and Privacy Preserving Keyword Searching for cloud storage services, Journal of Network and Computer Applications Vol. 35, pages: 927–933, Elsevier, 2012.
[6]. C. Orencik and E. Sava, Efficient and Secure Ranked Multi-Keyword Search on Encrypted Cloud Data, Proceedings in EDBT- ICDT, pages:186 -195, ACM: New York, 2012.
[7]. C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, Secure Ranked Keyword Search Over Encrypted Cloud Data. In ICDCS'10, pages 253- 262, 2010.
[8]. W. Zhou, L.Liu, H.Jing, C.Zhang, S.Yao, S.Wang, K-Gram Based Fuzzy Keyword Search over Encrypted Cloud Computing, Journal of Software Engineering & Applications, Vol. 6, pages: 29-32, 2013.
[9]. Mark D. Ryan, Cloud computing security: The scientific challenge, and a survey of solutions, Journal of Systems and Software Vol. 86, pages: 2263– 2268, Elsevier,2013.
[10]. http://en.wikipedia.org/wiki/Paillier_cryptosystem.
[11]. N. Islam, W. Puech, K. Hayat, R. Brouzet, Application of Homomorphism to Secure Image Sharing, Optics Communication Vol. 284, pages:4412-4429, Elsevier, 2011.

BIOGRAPHIE:

Received B.E degree in Computer Science and Engineering from Anna University of Technology, Tiruchirappalli, in 2012 She is Pursuing M.E degree in Computer Science and Engineering in R.M.K. Engineering College, Chennai. Her Research areas are focused on Cloud Computing and Network Security.