ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

Multi-Keyword Ranked Search and Dual Security on Cloud Data Using Homomorphic Encryption

Nikhitha K. Nair, Navin K.S.

PG Student, Dept. of C.SE., Sarabhai Institute of Science and Technology, Thiruvananthapuram, Kerala, India

Assistant Professor, Dept. of C.SE, L.B.S. College of Engineering, Poojapura, Thiruvananthapuram, Kerala, India

ABSTRACT The developments in the field of cloud computing increased the efficiency of data owners to deliver information technology services, where various resources such as data and software packages, which are stored in private servers can be retrieved from the internet with the help of various web-based tools and applications instead of having a direct connection with the individual servers. While accessing or sharing information stored in the cloud, the main concern is to enhance the security of data. Different encryption techniques can be used to provide strict privacy requirements. Dual encrypting the data using Advanced Encryption Standard and Homomorphic Encryption scheme enables the data users to ensure the integrity of requested data. Multi-keyword ranked searching enables the data users for efficient retrieval of searched query. Also Identity-Protocol performs the role of user identity verification providing identity tokens to the verified data owners in order to outsource their data in the cloud.

KEYWORDS: Cloud, Encryption, AES, XOR, Homomorphic Encryption

I. INTRODUCTION

Cloud computing is a model for delivering information technology services, in which various resources such as data and software packages which are stored in servers , are retrieved from the internet through web- based tools and applications ,rather than a direct connection to the server. However, cloud computing structure allows access to information as long as the electronic devices have access to the web. Cloud computing is so named because the information being accessed is stored in the "clouds", and does not require users to be in a specific place to gain access to the information stored in the clouds. Companies find that cloud computing reduce the cost of information management because, they does not require having their own servers and can use capacity leased from third parties. Also, cloud-like structure allows companies to upgrade software quickly.

Cloud computing brings on a new set of challenges as more owners of data are involved and their data are stored in offsite locations. When it comes to cloud computing, the security of personal information in the cloud is extremely important. The challenges include Encryption needs for cloud computing, Encrypting access to applications, Data ownership issues, Quality of service guarantees, Data retention issues, need for isolation management, Public cloud versus internal cloud security.

Considering the large number of data users and documents in the cloud, it is necessary to allow multi-keyword ranked search", in which multiple keywords can be given in search query, through which an effective search experience is felt by clients. It supports ranked response by a ranking algorithm that considers location and frequency of occurrence of the search word. Based on number of occurrences, relevant response is returned. That is, ranked results are given back to the clients; such that best result that well suits the search string is provided. For making multi-keyword ranked searching efficient, indexing is also performed. Here "Inverted sorted array indexing technique "is being used to retrieve the search results easily and efficiently.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

For protecting data privacy, encryption techniques can be performed by the data owners before outsourcing the data to public cloud. In "Single encryption", expert users can easily decrypt the data and reduce its essence. So, Dual encryption is required.

For security and privacy on data that is sent by the data owner to cloud for the data users, encryption that is performed is Dual encryption. The data sent by the data owner is encrypted and then stored in the cloud. But through this type of single encryption, expert network engineers can easily decrypt the data with assumption and reduce its essence. So, there is a need for dual-encryption to be provided for the data with pair key for a single data. These key can be hold by Cloud and the File Owner.

When the data owner sent encrypted data along with the index key to the cloud for the data users, then dual- encryption is performed on the data by the cloud, so that even the expert cannot decrypt the data easily. The data that is sent by the data owner to the cloud is first encrypted using AES (Advanced Encryption Standard) .The Cloud then performs Homomorphic encryption on the encrypted data sent by the data owner that generated homomorphic tags that indicates the correctness of data stored in the cloud. The data user then should perform dual decryption in order to retrieve the original data stored in the cloud. Here key generation is through "Attribute-based key generation technique".

System uses "Identity Protocol "that issue identity tokens to data owners based on their identity attributes for user account verification. For data owners to upload their encrypted data in the cloud, Identity protocol should grant permission to them. Identity protocol performs both user verification and file verification before uploading the encrypted files by the data owner in the cloud. That is, for the data owners to upload their encrypted file in the cloud, Identity protocol should grant permission by verifying the user and the encrypted file.

Dual Security with fully homomorphic encryption aims to bring forth the following objective-Dual encryption to provide additional security on stored data using AES encryption standard and Homomorphic encryption standard.

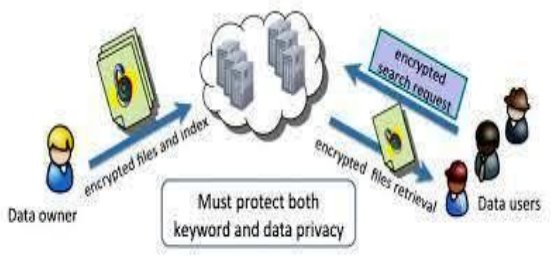

Fig.1. Architecture of the search over encrypted cloud data.

The remainder of this paper is organized as: The section II deals with the related work. In section III, the system model, design goals and the preliminary is introduced. Section IV deals with framework and privacy requirements for the entire system design, followed by section V, which describes the methodology of entire system functioning. Section VI describes the homomorphic encryption scheme used on the cloud data and we conclude the paper in Section VII.

II. RELATED WORK

In [2] attempt to solve the problem of supporting efficient ranked keyword search for achieving effective utilization of remotely stored encrypted data in Cloud Computing is performed. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy. To effectively support ranked search over encrypted file collection, the newly developed cryptographic primitive – order preserving symmetric encryption (OPSE) is used to achieve more practical performance. OPSE is a deterministic encryption scheme where the numerical ordering of plaintexts, gets preserved by the encryption function. Due to deterministic properties, if we use OPSE directly over these sampled relevance scores, the resulting cipher text

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

shall share exactly the same distribution as the relevance score. In [5] discusses the concept of Cloud computing to achieve a complete definition of what a Cloud is, using the main characteristics. It identifies different types of cloud systems and actors involved. Depending on the type of provided capability, there are three scenarios where Clouds are used: **1.** ***Infrastructure as a Service***, **2**. ***Platform as a Service*** and **3.** ***Software as a Service.*** In [8] a new framework is proposed for the problem of multi-keyword ranked search over encrypted cloud data, and to establish a variety of privacy requirements. Among various multi-keyword semantics, the efficient similarity measure is "coordinate matching", that is, as many matches are possible, to effectively capture the relevance of outsourced documents to the query keywords, and it uses "inner product similarity" to quantitatively evaluate such similarity measure. In [9] a viable solution for multi-keyword ranked query problems over encrypted data in the cloud environment is provided. A novel algorithm called MKQE uses a partitioned matrices approach. When the amount of encrypted data increases and more keywords need to be introduced, the searching infrastructure can be naturally expanded with the minimal overhead.

III. PROBLEM FORMULATION

A.SYSTEM MODEL

Considering a cloud data hosting service, involving three different entities playing different roles, as illustrated in figure.1: the data user, the data owner, and the cloud server. The data owner has a large collection of data documents to be outsourced to the cloud server. The data owner sends the encrypted data along with an index to the cloud server. The cloud server then performs dual encryption and stores that encrypted data in the cloud. Upon receiving the search request from the data users, the cloud server is responsible to search the index and return the corresponding set of encrypted data. For improving the accuracy of document retrieval, the search results should be ranked by the cloud server according to some ranking criteria such as coordinate matching. The data user has to perform dual- decryption in order to obtain the original data.

B. DESIGN GOALS

To enable dual security on cloud data, system design should achieve security guarantees as follows.

Multi-keyword Ranked Search:

For designing search schemes which allow multi-keyword query and provide result similarity ranking for effective data retrieval, instead of returning undifferentiated results.

Dual Security:

To prevent the cloud server from learning additional information from the data collection and the index, and to meet privacy requirements as specified in section IV.

A. PRELIMINARY ON ADVANCED ENCRYPTION STANDARD

When the data owner has a collection of data documents to be outsourced to the cloud server, the data owner performs encryption on the data to be outsourced using AES (Advanced Encryption standard) and generates a private key (owner key) using Attribute-based key generation technique.

D. PRELIMINARY ON HOMOMORPHIC ENCRYPTION STANDARD.

The cloud when receives the encrypted data of the data owner forwarded by the Identity protocol, the cloud performs dual-encryption using Homomorphic encryption scheme on them and generates a cloud key. Both the cloud key and owner key are required by the data users for decryption and retrieving the original data.

E. PRELIMINARY ON COORDINATE MATCHING.

Coordinate matching is a similarity measure which uses the number of query words appearing in the document to quantify the relevance of that document to the query. Because of huge amount of outsourced data, it is more flexible for the data users to specify a list of keywords indicating their interest and retrieve the most relevant documents with a rank order.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

F. PRELIMINARY ON XOR ENCRYPTION STANDARD.

For keyword privacy, the search query of data users are encrypted by them using XOR encryption standard. Upon receiving the search query from the data users, which is encrypted using XOR encryption scheme, the cloud server has the responsibility to search the index and return the corresponding set of encrypted documents. For improving the accuracy of document retrieval, the search results are to be ranked by the cloud server according to some ranking criteria such as coordinate matching techniques. Based on the frequency of occurrences of search keyword, relevant response is returned.

IV. FRAMEWORK AND PRIVACY REQUIREMENTS

In this section, the framework for dual security on encrypted cloud data are defined and established various strict system-wise privacy requirements for such a secure cloud data utilization system.

A. Framework

The entire system consists of five algorithms as follows.

Step 1: Setup: Data owner initializes the secret parameters of the data by creating key generation using Advanced Encryption Standard (AES).

Step 2: Build Index: The data owner builds a searchable index from unique words extracted from the data file collection. The owner then encrypts the data file collection and publishes the index, including keyword frequency, together with encrypted collection to the cloud server.

Step 3: Dual Encryption: Cloud server again encrypts the encrypted data and generates a cloud key using homomorphic encryption scheme.

Step 4: Trapdoor: To search the file collection for given keywords, the authorized data user generates and submits the search request in a secret form using XOR encryption scheme.

Step 5: Query Retrieval: Upon receiving the trapdoor, the cloud server will derive a list of matched file IDs and their corresponding relevance scores by searching the index via search index. The matched files will be sent back to the user in a ranked sequence based on relevance score.

B. Privacy Requirements

Privacy is the main factor when dealing with cloud data. As for data privacy, the data owner can use different encryption techniques to encrypt the data before outsourcing, and successfully prevent the cloud server from prying into the outsourced data. While dealing with index privacy, if the cloud server attempts to deduce any kind of association between the keywords and encrypted documents from the index, it may learn the major subject of a document. Therefore, the searchable index should be constructed to prevent the cloud server from performing such kind of association attacks. While dealing with keyword privacy, as data users normally prefer to keep their searching process from being exposed to others like the cloud servers, different encryption standard such as XOR can be employed on search query.

V. METHODOLOGY

Privacy-preserving Dual security on cloud data enhances security on data sent by the data owner to the cloud for storage. The entire system works as follows: When the data owner has a collection of data documents to be outsourced to the cloud server, the data owner performs encryption on the data to be outsourced using AES (Advanced Encryption standard) and generates a private key (owner key) using Attribute-based key generation technique.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

To enable effective searching process over the encrypted data collection, the data owner, before outsourcing the data, will first build a searchable index from the data collection and then outsource both the index and the encrypted document collection to the cloud server.

The Identity protocol then performs both data owner verification before allowing the data owner to outsource the data to the cloud. The identity protocol then grant identity tokens to the verified data owners which become their identity. Then the files of verified data owners are forwarded by the Identity protocol to the cloud for storage.

The cloud when receives the encrypted data of the data owner forwarded by the Identity protocol, the cloud performs dual-encryption using Homomorphic encryption scheme on them and generates a cloud key.

Upon receiving the search query from the data users, which is encrypted using XOR encryption scheme, the cloud server has the responsibility to search the index and return the corresponding set of encrypted documents. For improving the accuracy of document retrieval, the search results must be ranked by the cloud server according to some ranking criteria such as coordinate matching techniques. Based on the frequency of occurrences of search keyword, relevant response is returned.

The data users when receives the required files according to their search query, they must first decrypt the encrypted files (First decryption) using the cloud key and then again decrypt the files (Second decryption) using the owner key. Then only they can retrieve the original data sent by the data owner. There by dual security is provided.

VI. HOMOMORPHICENCRYPTION SCHEME

Homomorphic encryption is the encryption scheme which means the various operations on the encrypted data Homomorphic encryption is the encryption on the already encrypted data rather than the original data with providing the result as it is done on the plain text. The complex mathematical operations can be performed on the cipher text without changing the nature of the encryption. So, to allow the Cloud provider to perform the operations on encrypted data without decrypting them requires using the cryptosystems based on Homomorphic Encryption.

A. History of the Homomorphic encryption

In 1978 Ronald Rivest, Leonard Adleman and Michael Dertouzos suggested for the first time the concept of Homomorphic encryption. The encryption system of Shafi Goldwasser and Silvio Micali was proposed in 1982 was a provable security encryption scheme which reached a remarkable level of safety, it was an additive Homomorphic encryption, but it can encrypt only a single bit. In the same concept in 1999 Pascal Paillier was also proposed a provable security encryption system that was also an additive Homomorphic encryption. Few years later, in 2005, Dan Boneh, Eu-Jin Goh and KobiNissim invented a system of provable security encryption, with which we can perform an unlimited number of additions but only one multiplication. In 2009 Craig Gentry of IBM has proposed the first encryption system "fully homomorphic" that evaluates an arbitrary number of additions and multiplications and thus calculate any type of function on encrypted data.

B. Functions of Homomorphic Encryption Scheme.

The Homomorphic Encryption (H) scheme consists of four basic functions:
H = {Key Generation, Encryption, Decryption, Evaluation}

1. Key generation: The cloud will generate a cloud key on the encrypted data that is sent by the data owner to the cloud server.

2. Encryption: Using cloud key the cloud will again encrypt the data in the cloud that is sent by the data owner for storage.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

3. Evaluation: The cloud server has a function f for doing evaluation of cipher text and performed this as per the required function using cloud key.

4. Decryption: The data user decrypts the dual encrypted data using its private key and it gets the original result.

C. Properties of Homomorphic Encryption Scheme

The Homomorphic Encryption has mainly two properties,

Additive Homomorphic Encryption: A given cryptosystem is considered additively homomorphic iff
$\exists\Delta$: $\varepsilon(x1)\ \Delta\ \varepsilon(x2) = \varepsilon(x1 + x2)$
Where ε – encryption, x n- plaintext, cn – ciphertext and Δ- operation

Multiplicative Homomorphic Encryption:

A given cryptosystem is considered multiplicatively homomorphic iff

$$\exists\Delta: \varepsilon(x1)\ \Delta\ \varepsilon(x2) = \varepsilon(x1x2)$$

Where ε – encryption, x n- plaintext, cn – ciphertext and Δ- operation.

D. Applications of Homomorphic Encryption Scheme

The homomorphic encryption scheme can be used in applications where enhanced privacy is required such as banking transactions, voting systems and various cloud computing applications.

VI. CONCLUSION

This paper proposed dual encryption technique on cloud data to enhance security. The data send by the data owner is first encrypted using "AES (Advanced Encryption Standard)". The key generation is through Attribute based key generation technique. The encrypted data along with the index is uploaded to the cloud for storage. But before uploading, the Identity protocol should grant permission by verifying user identity and grant user with identity tokens. The data owner uploads the encrypted files to the cloud after generating signatures to recognize their blocks of files. The cloud then performs dual encryption on cloud data using homomorphic encryption scheme. The data users can download files by searching their request. For searching, keywords are also encrypted using "XOR encryption scheme" by the data users for keyword privacy. The cloud computing security based on Homomorphic encryption is a new concept of security which enables providing results of calculations on encrypted data without knowing the raw data on which the calculation was carried out, with respect of the data confidentiality.

REFERENCES

[1] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," *Proc. IEEE INFOCOM*, pp. 829-837, Apr, 2014.
[2] Cong Wang, Ning Cao, KuiRen and Wenjing Lou," Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE Transactions on Parallel and Distributes Systems, Vol.23, NO.8, AUGUST 2013.
[3] Craig Gentry, A Fully Homomorphic Encryption Scheme, 2009.
[4] ShashankBajpai and PadmijaSrivastava," A Fully Homomorphic Encryption Implementation on Cloud Computing", International Journal of Information & Computation Technology.ISSN 0974-2239 Volume 4, Number 8 (2014), pp. 811-816
[5] L.M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A Break in the Clouds: Towards a Cloud Definition," *ACM SIGCOMM Comput. Commun.Rev.*, vol. 39, no. 1, pp. 50-55, 2009.
[6] S. Kamara and K. Lauter,"Cryptographic cloud storage," in *RLCPS, January 2010, LNCS.Springer, Heidelberg.*
[7] S. Usha,Dr. A. Tamilarasi and R. Vijayakumar," Support Ranked Keyword on Remote Encrypted Data in Cloud", (International Journal of Engineering Trends and Technology (IJETT) – Volume 4 Issue 5- May 2013).
[8] AnkathaSamuyelu Raja and Vasanthi A, "Secured Multi-keyword Ranked Search over Encrypted Cloud Data ", International Journal of Advanced Research in Computer Science andSoftware Engineering-Volume 2, Issue 10, October 2012.
[9] ZhiyongXu, Wansheng Kang, Ruixuan Li, KinChoong Yow, and Cheng-ZhongXu,"Efficient Multi-Keyword Ranked Query on Encrypted Data in the Cloud", IEEE 18th International Conference on Parallel and Distributed Systems