Proxy Re-Encryption in Big Data Context

C. Selvi[1]
[1]. Associate Professor, CSE,
Velalar College of Engineering and Technology,

B. Akshata Kakra[2], A. Giridarani[3] ,V. Mithra[4]
[2, 3, 4]- Student CSE,
Velalar College of Engineering and Technology,
Erode-638012

***Abstract*:- With the growing amount of data, the demand of big data storage significantly increases. Through the cloud center, data providers can conveniently share data stored in the center with others. However, one practically important problem in big data storage is privacy. During the sharing process, data is encrypted to be confidential and anonymous. Such operation can protect privacy from being leaked out. To satisfy the practical conditions, data transmitting with multi receivers is also considered. Furthermore, this work proposes the notion of pre- authentication for the first time, i.e., only users with certain attributes that have already. The pre-authentication mechanism combines the advantages of proxy conditional re-encryption multi-sharing mechanism with the attribute-based authentication technique, thus achieving attributes authentication before re-encryption, and ensuring the security of the attributes and data. Moreover, this work finally proves that the system is secure and the proposed pre- authentication mechanism could significantly enhance the system security level.**

Keywords: Privacy-preserving, pre-authentication, proxy re-encryption, big data.

I. INTRODUCTION:

Nowadays, big data is a hot research topic. More and more users prefer to save their data in the cloud center because the cloud has a considerable amount of storage space, and users can download their data anywhere and anytime. People take photos, record music and do many other.

Operations on their personal equipments, producing large amount of data. As a matter of fact, the demand of cloud storage space is growing faster than ever before.

When people upload their data to the cloud, the first thing they may consider is whether the cloud storage is secure or not. They do not want other persons to peep their data without their permission. Public Key encryption is a mechanism designed for data providers to encrypt their data, and thus protecting the privacy of their data. Except the data receivers who have valid private key, no one can access the data. For example, in a hospital system, the patient records are too large and hard to store. A solution is to upload the massive data to the cloud for storage. Since everyone has access to the cloud, the data needs to be encrypted to prevent the private information of patients from being leaked out. When doctors intend to access the records, they decrypt the cipher text by using their keys and obtain the message they need. With this method of Public Key Encryption (PKE), the privacy required by the patients could be ensured.

Up to now, many cryptographic encryptions methods have been proposed to satisfy the requirements of privacy preserving in big data storage. However, most encryption methods such as the public key encryption are not anonymous, i.e., if the adversaries obtain the cipher texts, they can easily know the owner of the cipher text as well as who will receive the cipher text. The PKE cannot achieve the anonymity of the users send and receive the cipher text, so personal information may be leaked. If an adversary is able to achieve the cipher text, he can know whose key the cipher text is encrypted under, thus knowing the owner of the cipher text. To overcome this point, some anonymous encryption mechanisms have been proposed, e.g., anonymous mechanism. They achieve anonymity by removing the linkage between the data and the identity. Identities are splitted into two randomized complementary components and hide the identities of the receivers behind some randomization. Moreover, when users intend to store data in another cloud center, the data needs to be converted so as to be shareable among different cloud centers. Therefore, data receivers need to be updated. When users intend to share their data conditionally, like some parts of the data, the public key encryption cannot satisfy the users' requirements because when receivers know the key and decrypt the cipher text, they can achieve all the data. For example, when one user wants to share data about "music", he cannot do it because there are only Boolean conditions: receivers know everything or they know noting.

To solve the above problems, a great deal of effort has been made by the research communities. Boyen et al. proposed an identity-based encryption technique which is anonymous. By applying the method, linkage between users and cipher texts can be protected. Encrypt patients' PHRs to ensure that data store in the central will not be leaked out, which are similar to the problem studied in this work. However, many aspects have not been considered yet. For example, when a user transfers his/her data to another cloud center, a simple method is that the data provider decrypts the cipher texts first, and encrypts again before uploading it to another cloud center.

This method is not feasible when it comes to big data because time and computation cost will be enormous. Another drawback of this method is that when the provider is offline, no one can achieve the data. Entrusting the decrypt-encrypt-transmit task to a trusted third party is a good solution. But the information of receivers will be exposed to the third party during the re-encryption process. Blaze et al. proposed a technique named proxy reencryption. By applying a semi-trusted proxy and re-encrypt the cipher text, data can be shared without exposing information to the third party.

Furthermore, Green et al. proposed a technique called identity-based proxy re-encryption. They achieved control of access in storage in the network. Encryption techniques should be developed to meet requirements of the cloud users. Presented a proxy re-encryption technique called advanced multi hop- identity based conditional proxy re- encryption (AMH-IBCPRE) to realize receiver update and conditional share in big data storage.

In this work, we consider another situation. Users of the cloud center have the access to decide who to share with the data. They may intend to share data only with receivers who have certain attributes. Data providers and receivers have to verify the authenticity of each other to make sure that data and the identity won't be leaked out. The attributes also need to be protected. Proposed a verification mechanism to verify the authenticity of users' attributes. By applying this technique, we propose a mechanism called pre-authentication approach to proxy re-encryption. In our scheme, data providers can verify the authenticity of receivers. Once receivers' attributes do not meet the conditions, provider will not communicate with him anymore and he cannot obtain the data as well. This scheme is able to achieve four-fold properties: providing anonymity; updating receivers; verifying the attributes; sharing

conditionally. The main contributions can be concluded as follows: Propose a

new notion called pre- authentication mechanism in the model of MH-IBCPRE. Different from the existing work, the proposed mechanism can verify users' attributes before data sharing, thus satisfying the actual needs of users. The data of users can be shared with users having appointed attributes, and others have no access to the data.

proposed pre-authentication mechanism can provide multi-dimension privacy protection including data, user identities, and attributes. Only users whose attributes are authenticated could be qualified to share the data. This enhances the protection of user privacy.

We provide rigorous analysis to prove that our system is secure against chosen cipher text attacks, tracing attacks and collusion attacks. Side-by-side comparisons reveal that our scheme strengthens the security than MH-IBCPRE.

The remainder of the work is organized as follows. In, we describe the related works. After that, we give the system model and definitions. In, we introduce some preliminaries. Thus presented the complete system construction. Then, analyze the security of the system. Finally, conclude this work.

II. LITERATURE REVIEW:

Securing Personal Health Records in Cloud Computing:

Patient-Centric and Fine-Grained Data Access Control in Multi-owner Settings. Online personal health record (PHR) enables patients to manage their own medical records in a centralized way, which greatly facilitates the storage, access and sharing of personal health data. However, by storing PHRs in the cloud, the patients lose physical control to their personal health data, which makes it necessary for each patient to encrypt her PHR data before uploading to the cloud servers. Under encryption, it is challenging to achieve fine-grained access control to PHR data in a scalable and efficient way. For each patient, the PHR data should be encrypted so that it is scalable with the number of users having access. Also, since there are multiple owners (patients) in a PHR system and every owner would encrypt her PHR files using a different set of cryptographic keys, it is important to reduce the key distribution complexity in such multi-owner settings.

In this work, we propose a novel framework for access control to PHRs within cloud computing environment. To reduce the key distribution complexity, we divide the system into multiple security domains, where each domain manages only a subset of the users. In this way, each patient has full control over her own privacy, and the key management complexity is reduced dramatically. Our proposed scheme is also flexible, in that it supports efficient and on- demand revocation of user access rights, and break-glass access under emergency scenarios.

In recent years, personal health record (PHR) has emerged as a patient-centric model of health information exchange. A PHR service allows a patient to create, manage, and control her personal health data in a centralized place through the web, from anywhere and at any time which has made the storage, retrieval, and sharing of the medical information more efficient. In this way, the accuracy and quality of care are improved, while the healthcare cost is lowered.

Privacy-Preserving Cipher text Multi- Sharing Control for Big Data Storage:

In this work the need of secure big data storage service is more desirable than ever to data. The basic requirement of the service is to guarantee the confidentiality of the data. Moreover, the service also should provide practical and fine-grained encrypted data sharing such that a data owner is allowed to share a cipher text of data among others under some specified conditions. This work, for the first time, proposes a privacy- preserving cipher text multi-sharing mechanism to achieve the above properties. It combines the merits of proxy re- encryption with anonymous technique in which a cipher text can be securely and conditionally shared multiple times without leaking both the knowledge of underlying message and the identity information of cipher text senders/recipients. Security is the most important concern for any type of services which provides storage for data. Due to its efficient data processing capability cloud play a vital role in keeping big data. Many individuals and organizations can view, modify and update their data stored in the cloud through remote accessing. As increase in number of individual users and public and private organizations choose to upload their data in cloud force us to keep the data more securable from being hacked. The data of an individual user should be kept confidential and it should be accessed only by the authenticated users. While providing security, the most important aspect to be considered before storing the data is that, the anonymity of the service providers.

The services which are used for data storage should provide a high quality encrypted data sharing. The features mentioned above are commonly required to maintain secure processing, and these features are achieved by employing a new technique called cipher text multi sharing mechanism. It also ensures that, original message and information identity of cipher text senders and receivers is not leaked and it also ensures it is not vulnerable to cipher text attacks.

Attack Distribution and Distributed Forensics in Machine-to-Machine Networks:

In this work the advanced idea of machine- to-machine technology has attracted a new period of network revolution, evolving into a method to monitor and control global industrial user assets, machines, and the production process. M2M networks are considered to be the intelligent connection and communication between machines. However, the security issues have been further amplified with the development of M2M networks. It contains two modules: the attack detection module and the forensics analysis module. Machine-to-machine (M2M) networks are multidimensional networks that can use the Internet to achieve intelligent interactions between different machine terminals. They are composed of front-end sensors, equipment, transmission links, as well as back-end systems. First, M2M front-end nodes are responsible for collecting data and transmitting the data backhaul to the back- end systems. Second, the network transmission carrier is responsible for exchanging information between front-end and back-end sensors. Finally, the back-end control system can collect information to respond to the back-end nodes. M2M networks are conceived to communicate and connect between people, machines, and systems. They are associated with a large number of terminals. These terminals are vulnerable to attack, because they communicate through the wireless communication link integrated with different network protocols. Once M2M networks are attacked, the associated terminals will definitely be affected, and the whole network will be paralyzed. The network security issues are attracting much concern with the increasing development of M2M networks.

III. SYSTEM ANALYSIS: EXISTING SYSTEM:

In Existing system Public Key Encryption (PKE), the privacy required by the patients could be ensured. Up to now, many cryptographic encryptions methods have been proposed to satisfy the requirements of privacy- preserving in big data storage. However, most encryption methods such as the public key encryption are not anonymous, i.e., if the adversaries obtain the cipher texts, they can easily know the owner of the cipher text as well as who will receive the cipher text. The PKE cannot achieve the anonymity of the users send and receive the cipher text, so personal information may be leaked. If an adversary is able to achieve the cipher text, he can know whose key the cipher text is encrypted under, thus knowing the owner of the cipher text.

We propose a new notion called pre- authentication mechanism in the model of MH-IBCPRE. Different from the existing work, the proposed mechanism can verify users' attributes before data sharing, thus satisfying the actual needs of users. The data of users can be shared with users having appointed attributes, and others have no access to the data. Some existing work on multi sharing mechanism in big data. The comparison between our work and previous ones is shown afterwards.

DRAWBACKS:

- Encryption methods such as the public key encryption are not anonymous.
- The adversaries obtain the cipher texts, they can easily know the owner of the cipher text as well as who will receive the cipher text.
- They may intend to share data only with receivers who have certain attributes. Data providers and receivers have to verify the authenticity of each other to make sure that data and the identity won't be leaked out.
- The attributes also need to be protected. Proposed a verification mechanism to verify the authenticity of users' attributes.

PROPOSED SYSTEM:

The information of receivers will be exposed to the third party during the re- encryption process. Proposed a technique named proxy re encryption. By applying a semi-trusted proxy and re-encrypt the cipher text, data can be shared without exposing information to the third party. Furthermore proposed a technique called identity-based proxy re-encryption. They achieved control of access in storage in the network.

This work proposes the notion of pre- authentication for the first time, i.e., only users with certain attributes that have already. The pre-authentication mechanism combines the advantages of proxy conditional re-encryption multi-sharing mechanism with the attribute-based authentication technique, thus achieving attributes authentication before re- encryption, and ensuring the security of the attributes and data. Moreover, this work finally proves that the system is secure and the proposed pre-authentication mechanism could significantly enhance the system security level.

ADVANTAGES:

- First proposed the concept of the privacy of the keys is highly secure.
- The pre-authentication mechanism combines the advantages of proxy conditional re-encryption multi- sharing mechanism with the attribute-based authentication technique.

- Achieving attributes authentication before re-encryption, and ensuring the security of the attributes and data.
- Receivers who are qualified to know the data can use their keys to decrypt the cipher text, but others cannot, so data providers' privacy can be protected.
- To perfect the existing PRE system considered the scenario that data providers may want the data to be conditionally shared. That means receivers just obtain a part of the data instead of the whole. Such an assumption is more close to the reality.

IV. MODULE DESCRIPTION:

In this section, we present our whole system, including system set up, key generation, encryption, re-encryption, preauthentication and decryption. First, the parameters are set up and the secret keys are generated. The data is encrypted into cipher text. Then the generation of the re-encryption keys is carried on. After that, the attributes of the data receivers are verified, and only receivers with specific attributes have access to the re-encryption keys and re-encrypt the cipher texts. Finally, the decryption of the re-encrypted cipher texts is given.

Privacy-preserving

In this module PKE cannot achieve the anonymity of the users send and receive the cipher text, so personal information may be leaked. If an adversary is able to achieve the cipher text, he can know whose key the cipher text is encrypted under, thus knowing the owner of the cipher text. To overcome this point, some anonymous encryption mechanisms have been proposed, e.g., anonymous mechanism. They achieve anonymity by removing the linkage between the data and the identity. Identities are splitted into two randomized complementary components and hide the identities of the receivers behind some randomization.

Pre-authentication

This module is may intend to share data only with receivers who have certain attributes. Data providers and receivers have to verify the authenticity of each other to make sure that data and the identity won't be leaked out. The attributes also need to be protected. Proposed a verification mechanism to verify the authenticity of users' attributes. By applying this technique, we propose a mechanism called pre-authentication approach to proxy re-encryption. In our scheme, data providers can verify the authenticity of receivers. Once receivers' attributes do not meet the conditions, provider will not communicate with him anymore and he cannot obtain the data as well.

Proxy re-encryption

This module is help to applying a semi- trusted proxy and re-encrypts the cipher text; data can be shared without exposing information to the third party. Furthermore, Green et al. proposed a technique called identity-based proxy re-encryption. They achieved control of access in storage in the network. Encryption techniques should be developed to meet requirements of the cloud users. Presented a proxy re-encryption technique called advanced multi hop- identity based conditional proxy re- encryption (AMH-IBCPRE) to realize receiver update and conditional share in big data storage.

Big data key generation

This module is entrusting the decrypt- encrypt-transmit task to a trusted third party is a good solution. But the information of receivers will be exposed to the third party during the re-encryption process. Proposed a technique named proxy reencryption. By applying a semi-trusted proxy and re- encrypt the cipher text, data can be shared without exposing information to the third party. Furthermore proposed a technique called identity-based proxy re-encryption. But the information of receivers will be exposed to the third party during the re- encryption process. Proposed a technique named proxy reencryption. By applying a semi-trusted proxy and re- encrypt the cipher text, data can be shared without exposing information to the third party. Furthermore proposed a technique called identity-based proxy re-encryption. They achieved control of access in storage in the network.

SYSTEM REQUIREMENTS:

HARDWARE REQUIREMENTS:

System	: Pentium core
Hard Disk	: 80 GB.
RAM	: 1 GB DDR2
Key Board	: LG
Mouse	: Logitech.
Monitor	: 15 inch VGA

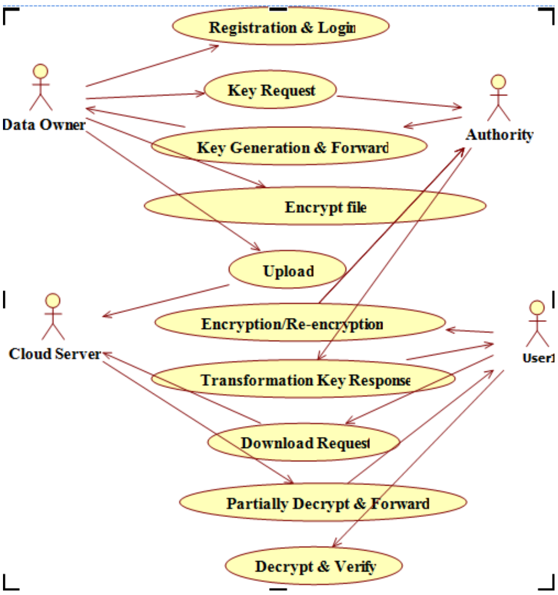
SOFTWARE REQUIREMENTS:

Operating system	: Windows 7 Front
End	: JDK 1.7/Net Beans

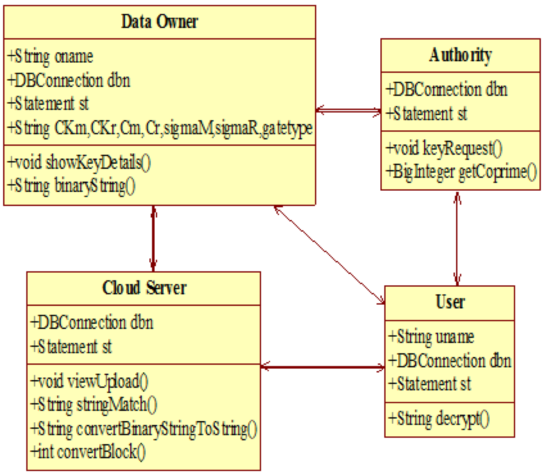
8.0
Coding Language : Java

UML DIAGRAM:

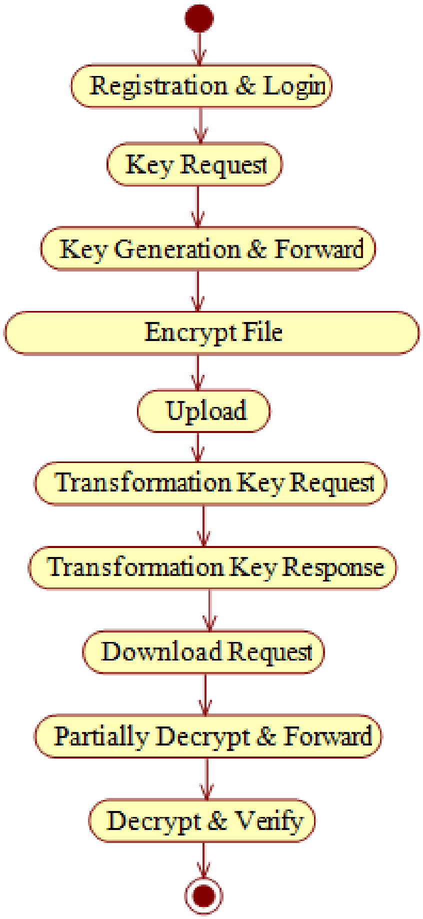
Use Case Diagram:



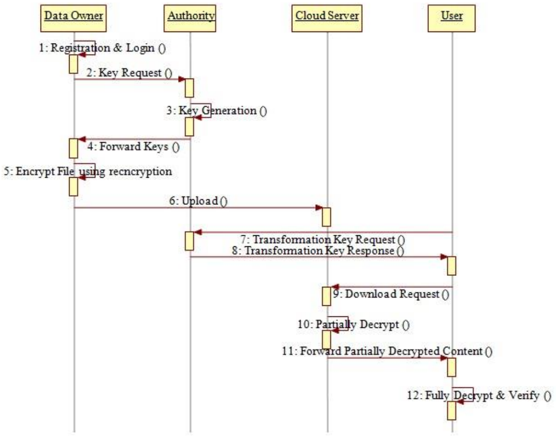
Class Diagram:



Activity Diagram:



Sequence Diagram:


Data Owner
Authority
Cloud Server
User
1: Registration & Login ()
2: Key Request ()
3: Key Generation ()
4: Forward Keys ()
5: Encrypt File using recncryption
6: Upload ()
7: Transformation Key Request ()
8: Transformation Key Response ()
9: Download Request ()
10: Partially Decrypt ()
11: Forward Partially Decrypted Content ()
12: Fully Decrypt & Verify ()

EXPERIMENTAL SETUP:

Comparing results when data are on disk versus in cache shows that disk throughput bounds IB-DPDP's performance when accessing all blocks. With the exception of the first blocks of a file, I/O and the challenge computation occur in parallel. Thus, Proxy-Conditional-Re-Encryption generates proofs faster than the disk can deliver data: 1.0 second versus 1.8 seconds for a 64 MB file. Because I/O bounds performance, no protocol can outperform Proxy-Conditional-Re-Encryption by more than the startup costs. While faster, multiple- disk storage may remove the I/O bound today. Over time increases in processor speeds will exceed those of disk bandwidth and the I/O bound will hold. Sampling breaks the linear scaling relationship between time to generate a proof of data possession and the size of the file. At 99% confidence, Proxy-Conditional-Re Encryption can build a proof of possession for any file, up to 64 MB in size in about 0.4 seconds. Disk I/O incurs about 0.04 seconds of additional runtime for larger file sizes over the in-memory results. Sampling performance characterizes the benefits of IB-DPDP. Probabilistic guarantees make it practical to use public-key cryptography constructs to verify possession of very large data sets. Table 1 and 2 shows the preprocessing accuracy and overall accuracy of the proposed and existing system.

Algorithm	Time in ms	File size in kb
Existing	4.5	2.5
Proposed	4.0	2.5

Table: 1 Preprocessing accuracy comparison between existing and proposed work

Algorithm	Over all Accuracy in percentage	Data integrity per block(for 100 percentage)
Existing	78	93
Proposed	83	98

Table: 2 Overall Accuracy and Data integrity comparison between E-IBE with Proposed Proxy-Re-Encryption approach.

V. CONCLUSION & FUTUREWORK:

In this work, we realize multi-sharing, anonymous and CCA-secure data sharing in big data context. Furthermore, we propose a new notion called pre-authentication in the proxy re-encryption system, which can ensure that only users whose attributes have been verified are permitted to obtain the data and provide well protection for the private attributes.

The pre-authentication function greatly facilitates the needs of the users. Besides, we prove that users' data, identities and attributes are protected, and the pre-authentication process enhances the security of the system. To the best of our knowledge, we are the first to propose the concept of preauthentication in this aspect.

REFERENCES

1. X. Boyen and B. Waters, **"Anonymous hierarchical identity-based encryption (without random oracles (lecture notes in computer science),"** Advances in Cryptology, vol. 4117, pp. 290–307, Aug 2006.
2. K. R. M. Li, S. Yu and W. Lou, **"Securing personal health records in cloud computing: Patient- centric and fine-grained data access control in multi-owner settings,"** Security and Privacy in Communication Networks -, International ICST Conference, SECURECOMM, pp. 89–106, 2010.
3. E. H. J. Benaloh, M. Chase and K. Lauter, **"Patient controlled encryption: Ensuring privacy of electronic medical records,"** ACM Cloud Computing Security Workshop, pp. 103–114, 2009.
4. M. Green and G. Ateniese**, "Identity-based proxy re-encryption,"** Applied Cryptography and Network Security (Lecture Notes in Computer Science), vol. 4521, pp. 288–306, 2007.
5. W. S. K. Liang and J. Liu, **"Privacy- preserving cipher text multisharing control for big data storage,"** IEEE Transaction on Information Forensics and Security, vol. 10, no. 8, Aug 2015.
6. J. S. L. Guo, C. Zhang and Y. Fang, **"A privacy-preserving attribute based authentication system for mobile health networks,"** IEEE Transaction on Mobile Computing, vol. 13, no. 9, Sep 2014.
7. K. Wang, Y. Shao, L. Shu, G. Han, and C. Zhu, **"Ldpa: A local data processing architecture in ambient assisted living communications,"**IEEE Communications Magazine, vol. 53, no. 1, pp. 56–63, Jan 2015.
8. K. Wang, Y. Shao, L. Shu, Y. Zhang, and C. Zhu, **"Mobile big data fault-tolerant processing for ehealth networks,"** IEEE Network, vol. 30, no. 1, pp. 1–7, Jan 2017.
9. K. Wang, M. Du, D. Yang, C. Zhu, J. Shen, and Y. Zhang, **"Game theory-based active defense for intrusion detection in cyber physical embedded systems,"** ACM Transactions on Embedded Computing Systems, vol. 16, no. 1, Article 18, Oct 2016.
10. P. L. J. Shao and Y. Zhou, **"Achieving key privacy without losing coca security in proxy re- encryption,"** Journal of Systems and Software, vol. 85, no. 3, pp. 655– 665, 2011.