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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

# A Survey on Encrypted File Sharing and Anomaly Detection in Cloud Computing

Miss. Neha Amol Kale

Assistant Professor, Department of Computing Science and Engineering, Sandip University, Nashik, India

**ABSTRACT:** Cloud Computing is the emerging technology that combines the concept of "Software-as-a-Service" and "Utility Computing", provides the on-demand services to the end users. Security is the important aspect and has various issues and problem in cloud computing. Nowadays many organizations are moving their data on the cloud, by using File Syncing and Sharing Services. End users uses their own devices to access the data and due to this there is rise in the new challenge for preventing the player/decoder abuse. In this paper, a system is developed called as PHE that is Partially-Ordered Hierarchical Encryption which implements partial order key hierarchy. Partial order key hierarchy is same as role hierarchy used in Hierarchical Role Based Access Control(HRBAC). This paper also introduces anomaly detection by using audit, pattern matching and risk assessment. This anomaly detection will identify the suspected players and will trace and revoke the authorities of the suspected players.

**KEYWORDS**: Audit, Cloud, File Syncing and Sharing, Pattern Matching, Partial Ordered Hierarchy, Revocation, Risk Assessment, Role Hierarchy, Security, Traitor Tracing**.**

## I. INTRODUCTION

Cloud Computing is technology that provides different services to the users and it is used to manipulate, configure, and access the resources remotely. Cloud offers different services such as data storage, infrastructure, and application. We can access the data or services from any location at any point of time. There are different services of cloud that are available such as Box, Drop Box, Sky Drive, Sugar sync for individual and small to medium business. These cloud storages provide on demand capacity, low cost services and long-termarchives. The cloud storage and services make people life more convenient as user can access the data, application from anywhere at any time via any available device such as computers, mobile phones, etc. so many organizations and individuals have moved their personal data, large archive system into cloud storage. Cloud has become a necessity to many individual, organization and government use [1].

The cloud storage service is initiated by individual users who store data and download it to sync and collaborate. So, cloud-based storage provides file syncing and sharing services. In File Sharing, users can access the files that are located anywhere at any time from various end devices, and can also edit the files together. File Syncing is a backup mechanism that is used for syncing the data or information across various devices such as personal computer, smartphone etc [1].

Cloud is network of computers that are connected to each other in same or different geographical locations, that operates together to provide services to the customers that are having different need and workload on the demand basis. The services of cloud are provided to the users on the basis of usage that is user will pay as per the use and demand. These services are in the form of platform, Infrastructure or software [2]. The services on the cloud are provided by the cloud providers and by using these services, the cloud users build the applications and deliver them to the customer. So, the cloud users do not have to think about installation or maintenance of hardware and software they need. And these services are affordable to the users as they have to pay as per their usage. Instead of establishing IT infrastructure themselves.The cloud services help the cloud users to reduce their efforts and expenditures in the IT field.

Cloud computing model combines concept of software as a service(SaaS) and utility computing. It provides convenient and on demand services to the users. Cloud service providers provides different services to the end users.

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

The providers and the customers that uses the services have to make it sure that data or information is kept secure from the external entities or attacks so that the information does not get lost. [3].

In cloud computing security and privacy of the data is one of the major issues. The privacy of user's data needs to be preserved. Security is a problem that must be considered for deploying a file sharing and syncing service. Many surveys [4] have proved that cloud users always worry about the privacy and the security of their data and these two things are always the obstacle in cloud computing. The multi-tenant nature of cloud makes it vulnerable to data leakage, threats and malicious attacks. So there must some kind of policies that the enterprise must have such as Role Based Access Control or Attribute Based Access Control [5] so that the privacy and confidentiality can be preserved As the data on the cloud is sensitive it needs to be encrypted before storing it on the websites, but there is one drawback of encrypting data, the data will be shared at coarse grain level i.e. giving your private key toanother party. So, toovercome this problem, ABE with fine grain access control isintroduced, where cipher text is labelled with attributes and private keys are associated with access structure [6].

In Role Based Access Control(RBAC) every user is assigned with the roles that he wants to perform and by assigning these roles to the users, the unauthorized access will be restricted. It is also called as role-based security. This policy is defined around the roles and privileges. RBAC is used for administrating the security of various large organizations. The roles in RBAC represents the job function within the organization and authorization is granted to the role instead of single user. The authorization is restricted to data objects and resources needed by college. Users are authorized to play the appropriate role. This role in RBAC can support security policies of the organization [7]. Role hierarchy is the means where the structuring roles to reflect the structure of the organization. For this a partial order relation is applied on the roles known as role hierarchy [7].

## II. RELATED WORK

### 2.1 CP-ABE Traceability

2.1.1 Introduction

The Ciphertext Policy Attribute Based Encryption is a technique in which the ciphertext are associated with the access policies and the attributes are shared with each other. CP-ABE has the fine grain access control over the encrypted data. But there is a concern about the CP-ABE method, the malicious user can share the attribute with other users which cause the leakage of decryption privileges for financial gain. To overcome this problem a new technique is developed called Traceable CP-ABE. It consists of 2 levels, While box traceability and black box traceability. [8]

Whitebox traceability scheme may not able to find the malicious users that make the decryption black box. But if the decryption Blackbox is given and the tracing algorithm and decryption key is kept hidden then Blackbox traceability can find the malicious user. Blackbox traceability has 2 types: Key-like decryption black box, and Policy-Specific black box.

In Key-Like decryption black box the abnormal users use his decryption key is to build a decryption black box and sells it on eBay to get the financial profit. A law enforcement agency gets a warrant to search a suspect's computer and finds a decryption black box. As the suspect, might try to destroy evidence, the explicit description of the black box's ability might be gone, while the law enforcement agency only has certain clue on the certain access policy associated to the ciphertexts that the black box can decrypt. We consider this decryption black box as a policy-specific decryption black box.

2.1.2 System Architecture

The CP-ABE scheme consists of four algorithms [8]

- Setup: In this algorithm the input is some security parameter, no of users in system, and output is the public parameter and master secret key.
- KeyGen: In this algorithm the input is the public parameter, the master secret key and attribute set S and the output is the decryption key.
- Encrypt: This step takes the public parameter, a message M and some access policy as input and returns ciphertext as the output.

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

- Decrypt: Takes a input the public parameter, a cipher text and a decryption key and outputs the message M.

2.1.2 Pros

1. The Traceable CP-ABE scheme is proved to be fully secured, highly expressive that is it can support any monotonic access structure.
2. Fully Collusion resistant and efficient with sublinear overhead

2.1.3 Cons

1. This Study is based on composite order bilinear groups but it is inefficient when compared with the prime order bilinear groups.
2. Not adaptively traceable against the policy specific decryption black box

**2.2 Augmented CP-ABE**

2.2.1 Introduction

In this scheme the index hiding property is redefine and can be transformed into a traceable CP-ABE against policy specific decryption black box.[8]
The AugCP-ABE has four algorithms.
Setup: This algorithm takes the security parameter, no of users as input, runs in polynomial time and gives the public parameter and a master key as the output.
KeyGen: This algorithm takes as input the public parameter that is generated in the previous algorithm and a attribute set and gives the output a private key which has a unique index.
Encrypt: It takes input the public parameter, a message, access policy and index and gives the output a cipher text
Decrypt: Input for this algorithm is public parameter, cipher text, attribute set and returns the message as output.

**2.3Smart Grid Communication Infrastructure**

2.3.1 Introduction

The information about the user's actions, priority and beliefs can be extracted from high resolution load data or information. This information of user needs to be preserved, but this information is used in ways that may violate the user's privacy. User is having little control on the data. So, to increase the control of user over information, there is need that the load data be represented in multiple resolution and each resolution is secured with different key. To make it possible it is necessary to introduce a suitable hierarchical key management. Because of this the user has the freedom to decide who will access the data and at what granularity[9].

To preserve the data and make the communication secure, a system with confidentiality, authentication, integrity is needed within the smart grid. The communication between two parties must be secured and no other party will be able to decrypt the communication channel and if access is granted. Every entity will get the access of data or resources based on his needs. The data is stored as long as it's needed and the user is informed about the usage of the data[9].

There are various attacks on the smart grid infrastructure, which needs to be stopped. There are various techniques that are assumed to be safe.

2.3.2 System Architecture

The PKI is a key management system that manages a big infrastructure and this approach is based on certificate based Public Key Infrastructure. The PKI is managed by the grid operator it makes use of bridges to enable communication with other PKIs.

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

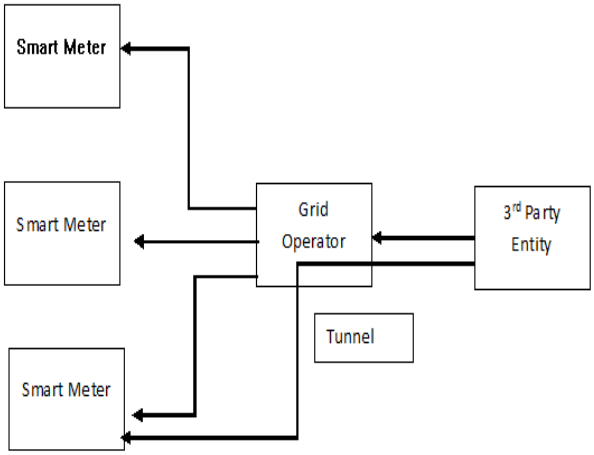


**Fig: Smart Grid**

The smart meter plays important role in the architecture and it is a trusted device. Smart meters generate keys and stored the key in such a way that it is not readable to others.[8] The Smart meter computes the cryptographic functions. As shown the meters are connected directly to the Grid operator. The third-party entity accesses the smart meter via API that is provided by the grid operator

2.3.3 Pros
There are two benefits of this approach
1. The grid operator chooses the technology of how to communicate with the smart meters.
2. As the smart meters are not exposed directly to the third party, the security is increase by allowing only the authorized users to communicate with smart meter.
3.Monitoring and blocking unauthorized traffic by grid operator.

### 2.4 Attribute Based Encryption

2.4.1 Introduction
Attribute Based Encryption(ABE) is a scheme in which users attribute set is used for generation of secret key and it uses access structure that control the access to the data. ABE scheme is useful for secret sharing purpose among the groups. ABE scheme supports ciphertext policy access control and key policy access control.ABE scheme is derived from the Identity Based Encryption and there are certain fundamentals that empower the ABE scheme such as bilinear maps, lattice approach.

1.Bilinear Maps
Bilinear maps solve the practical issues such as computational requirements and is efficient at key generation, precise and secure. Bilinear maps have some properties such as Bilinearity, Computability and Non-degeneracy. There is bilinear pairing in bilinear maps that is used for pairing the points from source group to target group. Bilinear pairing practically is the computation of public and secret key pairs.Bilinear maps is a solution to the IBE and ABE scheme. Bilinear maps solution is based on Computational Diffie Hellman(CDH).

2. Lattice Approach
The lattice approach is more efficient than bilinear maps. It only requires small number arithmetic and basic algebra for implementation.

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

2.4.2 Architecture

The ABE scheme is derived from IBE so both the algorithms are same. ABE scheme has four algorithms.

1.Setup: This algorithm defines the bilinear mapping and takes as input the security parameter, the size of attribute and returns the public key and master key as output.

2.KeyGen: Used for generating keys and takes the master key and access tree as the input and produces the secret key as input.

3.Encrypt: This algorithm is used encrypting the message M into the ciphertext

4.Decrypt: It takes the ciphertext and decrypt it into the original message M.

2.4.3 Pros

The ABE scheme has various features such as scalability, threshold policy access control, Key policy access control, ciphertext policy access control, and many more

**2.5 Traitor Tracing, and Revocation**

2.5.1 Introduction

Traitor tracing scheme is used to protect the broadcast encrypter. It has four algorithms. The First one is setup algorithm which is used for generation of the secret keys and public parameter for the user. The encrypt and decrypt algorithm is used for encrypting and decrypting the message and the trace algorithm is used for tracing the pirate decoder or the user who tries to decrypt the message illegally [10].

The revocation scheme provides the additional function of user revocation. The revocation algorithm is used to revoke the authorities of the unauthorized users who tried to decrypt the message or access the data illegally. There are different primitives that are sufficient to implement efficient traitor tracing and revocation system, such as Private Linear Broadcast Encryption (PLBE) and Augmented Broadcast Encryption (AugBE).

The PLBE scheme consists of four algorithms. Setup, Encrypt, Decrypt, and TrEncrypt. The Setup, Encrypt and Decrypt algorithms are similar to the other algorithm and TrEncrypt algorithm is used for tracing purpose. The AugBE consists of three algorithms, Setup, Encrypt and Decrypt. PLBE scheme is sufficient to construct a traitor tracing system and an AugBE scheme is sufficient to construct a trace & revoke system

2.5.2 Pros

This technique gives 10 times faster decryption, 6 times faster encryption, and reduces 50% of the ciphertext size.

## III. COMPARATIVE ANALYSIS

| Sr. No | Parameter | Traceable CP-ABE | AugCP-ABE | ABE | Tracing and Revocation | RBAC |
|---|---|---|---|---|---|---|
| 1 | **Access Control** | Monotonic access control | Monotonic access control | Non-monotonic policy, Key policy, Ciphertext policy, Threshold Policy | monotonic access structures | policy neutral access control |
| 2 | **Efficiency** | Inefficient when compared with prime order bilinear group | Inefficient when compared with prime order bilinear group | Efficient revocation as compared to previous ABE | Efficient | Efficient |
| 3 | **Performanc** | Good as | Good | Good | Good as | Good as lower overloads |

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

| | | | | | | |
|---|---|---|---|---|---|---|
| | **e** | compared to CP-ABE | | | encryption is 6 times faster, decryption is 10 times faster and ciphertext size is 50% less | for large-scale systems, Minimal administrator work |
| **4** | **Scalability** | Scalable as compared to CP-ABE | Scalable | Scalable in comparison with previous ABE schemes | Scalable | Scalable |
| **5** | **Collusion Resistant** | Yes | Yes | Yes | Yes | Yes |

## IV. PROPOSED SYSTEM

*A.System Architecture*

The main aim of the system is to provide security to the data stored on the cloud. Users store their confidential file or data on the cloud and access them through file syncing and sharing services(FSS). FSS allows users to sync and share the files and access them from anywhere at any time and from variety of devices and collaboratively edit them together. The existing system uses Role Based Access Control scheme which implements the partial order key hierarchy that is like to role hierarchy. This scheme provides security mechanism such as traitor tracing and revocation. The abnormal player/decoder which tries to access the files illegally are traced and their authorities are cancelled by using tracing and revocation algorithm respectively.

To improve the security and privacy of the outsourced data a more comprehensive anomaly detection is implemented by using techniques such as audit, pattern matching, risk assessment. The anomaly detection will recognize the unauthorized access that is the anomaly behaviour of the player against RBAC policies and the license terms. Anomaly detection also detects the non-normal distribution exceeding the pre-set bounds & identify the player abuse outside the permitted range.

A PHE cryptosystem is developed that consists of 6 tuples (S, J, E, D, T, R)

**1. Setup:** The first tuple is Setup. It takes the input parameters such as partial order hierarchy that is omega, security parameter s, maximum collusion number t. It outputs the encryption key that is required for encrypting the data, set of public parameter p and a master key that will be the manager secret.

**2. Join**: It includes two algorithms. one for joining the individual user and the other one for joining group of users.

**2.1. Join-user:** This tuple is used for joining individual user. It takes as input the master key and the identity of the user. It outputs the decryption key of the user and this key is sent to the user secretly.

**2.2. Join-Group:** This tuple is used for joining group of users. It takes as a input the master key and the group identifier and generates some public parameters and the decryption key of that class.

**3. Encrypt**: This tuple is used for encrypting the data stored on the cloud. It encrypts the message M and outputs the cipher text. This tuple uses various encryption algorithms such as AES or RSA

**4. Decrypt:** It decrypts the cipher text using the decryption key and outputs the message M.

**5. Trace:** It is used to trace the abnormal players and detect the traitor from the players. Tracing algorithm consists of 2 aspects.

**6. Revoke:** This tuple will cancel all the authorities of the abnormal players and revoke them. There is public revocation algorithm that will revoke the suspicious users or the classes.

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

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 3, March 2019

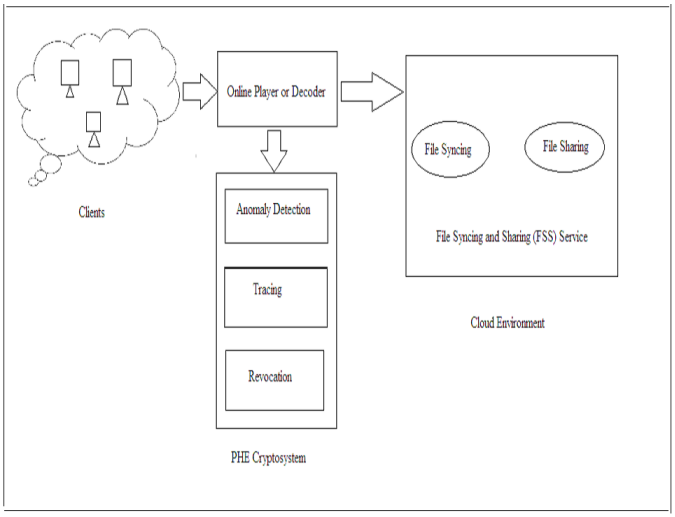


**Fig 1: System Overview**

*B. Working*

In the proposed system, the users are installed on client side. Users are the one who wants to access the files or the data stored on the cloud. Users can access the file from anywhere and from any device. The player/decoder are the one whose provides the interface to access the data on the cloud. In the file syncing and sharing service(FSS), File sharing allows the user to access the file from anywhere at any time and from any device and syncing is the backup mechanism for syncing the data across the multiple devices. The FSS service is on the cloud side. The FSS model provides prevention against player abuse and enhanced the protection from unauthorized access. This model uses Role based access control which is based on role hierarchy, and is recognized for its support for simplified administration and scalability to work in collaboration with teams.

When the user wants to access the data stored on the cloud, it tries to access the data through the player/decoder. The player provides the interface to the user but the behaviour of the player might be suspicious so anomaly detection is applied on the player's behaviour. After applying the anomaly, the suspected player is detected. Once the player is detected, the traitors are found out from that players. After finding out the traitors, they are traced out and revocation mechanism is carried out on the abnormal players.The traitor tracing and revocation algorithm is useful for tracing the abnormal players and revoking the authorities of that player. A group-oriented cryptosystem called PHE is applied to anomaly detection and traitor tracing and revocation.

## V. DISCUSSION AND CONCLUSION

I analysed various techniques of cloud security and identified the problem area where there is need a find a solution over a privacy and security of data on cloud using different methodology and features. In this paper, I have described the new system that will help in protecting the data against anomalies.

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

International Journal of Innovative Research in Computer and Communication Engineering

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

*Website: www.ijircce.com*

**Vol. 7, Issue 3, March 2019**

## REFERENCES

[1] Yan Zhu, Guohua Gan, Ruiqi Guo, and Dijiang Huang," PHE: An Efficient Traitor Tracing and Revocation for Encrypted File Syncing-and-Sharing in Cloud", IEEE Transaction on Cloud Computing,2016.
[2] Prince Jain," Security Issues and their Solution in Cloud Computing", International Journal of Computing Business Research,2012
[3] Santosh Kumar and R. H. Goudar, "Cloud Computing Research Issues, Challenges, Architecture, Platforms and Applications: A Survey", International Journal of Future Computer and Communication, Vol. 1, No. 4, December 2012.
[4]F. R. Institute, "Personal data in the cloud:A global survey of consumer attitudes", http://www.fujitsu.com/downloads/SOL/fai/reports/fujitsu/personaldata-in-the cloud.pdf, 2010.
[5] ZhiQiao, Shuwen Liang, Spencer Davis and Hai Jiang," Survey of Attribute Based Encryption", IEEE Conference,2014
[6] V. Goyal, O. Pandey, A. Sahai, and B. Waters," Attribute-based encryption for fine-grained access control of encrypted data", in ACM Conference on CCS,pp.8998,2006
[7] A. Fiat and M. Naor," Broadcast encryption", in Advances in Cryptology (CRYPTO93), vol. 773 of LNCS. springer-verlag, pp. 480491,1994
[8]Zhen Liu, Zhenfu Cao, "Traceable CP-ABE: How to Trace Decryption Devices Found in the Wild", IEEE Transaction on Information Forensics and Security, Vol. 10, N0. 1, January 2015
[9] Christian D.Peer, Dominik Engel, Stephen B.Wicker, "Hierarchical Key Management for multi-resolution Load Data Representation", IEEE International Conference on Smart Grid Communications, 2014
[10] D. Boneh and B. Waters," A fully collusion resistant broadcast, trace, and revoke system", in ACM Conference on Computer and Communications Security, pp. 211220,2006.
[11]David F. Ferraiolo, Janet A. Cugini, D. Richard Kuhn, "Role-Based Access Control (RBAC): Features and Motivations"
[12] H. Chung, J. Park, S. Lee, and C. Kang," Digital forensic investigation of cloud storage services", Digital Investigation, vol. 9, no. 2, pp. 8195, 2012.
[13] Chen, S. Nyemba, and B. Malin," Detecting anomalous insiders in collaborative information systems", Dependable and Secure Computing, IEEE Transactions on, vol. 9, no. 3, pp. 332344, May 2012
[14] M. Blanton and K. B. Frikken, "Efficient Multi-dimensional key management in broadcast services", in ESORICS, pp. 424 440,2010
[15] S. Garg, A. Kumarasubramanian, A. Sahai, and B. Waters," Building efficient fully collusion-resilient traitor tracing and revocation schemes", in Proceedings of the 17th ACM Conference on Computer and Communications Security, CCS 2010, Chicago, Illinois, USA, October 4-8, pp. 121130,2010.
[16] M. J. Atallah, M. Blanton, N. Fazio, and K. B. Frikken," Dynamic and efficient key management for access hierarchies", ACM Trans. Inf. Syst. Secur., vol. 12, no. 3, 2009.
[17] N. Attrapadung and H. Imai," Conjunctive broadcast and attribute-based encryption", in Pairing-Based Cryptography - Pairing 2009, Third International Conference, Palo Alto, CA, USA, August 12-14, 2009, Proceedings, pp.248265,2009
[18] Y. Chung, H. Lee, F. Lai, and T. Chen, "Access control in user hierarchy based on elliptic curve cryptosystem", Information Sciences, vol. 178, pp. 230243,2008.
[19] E. Bertino, N. Shang, and S. Wagstaff," An efficient time-bound hierarchical key management scheme for secure broadcasting", IEEE Trans. on Dependable and Secure Computing, vol. 5, no. 2, pp. 6570, 2008
[20] R. Ostrovsky, A. Sahai, and B.Waters, Attribute-based encryption with nonmonotonic access structures", in ACM Conference on Computer and Communications Security, pp. 195203,2007.