Fine-grained Access Control Scheme Based on Cloud Storage

Xiaojie Niu
Deptment of IT Services,
Computer Application Institute of Nuclear Industry
Beijing,China
E-Mail:shybull@163.com

***Abstract*—Cloud storage originates from cloud computing, and brings new model of data sharing and storage, which provided great convenience for users. However, cloud environment confronts with many problems, and one of the most important problems is the security problem. This paper researched on security problem based on encryption, and proposed a secure and efficient scheme according to system characteristics of data storage on cloud platform, and applied it in cloud storage system with fine-grained access control based on CP-ABE. Through the analysis and comparison, the experimental results showed that the proposed scheme optimized the user revocation, reduced the time of data owner to manage data, and realized the safe sharing and efficient storage of sensitive data in the public cloud storage. Finally, the technology in this paper was an optimization both on security and the whole expense, and it must have a good prospect in the future.**

Keywords-access control; cloud computing; attribute encryption algorithm; CP-ABE encryption; cloud storage.

I. INTRODUCTION

According to the security problem of cloud storage, users need to encrypt data, and in the cloud storage system to achieve access control. In the current storage system, security and performance are always opposite. When the introduction of security means that will spend more time. However, you can make a balance between the security of the system and the overall overhead by reducing the storage space. So this paper proposes a cloud storage system with high efficient user revocation based on CP-ABE[1].

In the 90s of last century, the network as a novel and convenient information media, gradually being recognized. People realize that it has a huge scale of computing resources, fascinated by its huge application prospects, and to study how to use these resources efficiently and easily. With the popularization and application of cloud computing technology, people have the ability to use large-scale distributed computing resources in the network. Cloud computing as a hot topic of research and application in recent years, most IT companies and industry insiders believe that the next generation of computer network application technology core architecture. Under the cloud computing environment, users do not have to spend the high cost of hardware and software to powerful computing resources and huge storage capacity, all of which can be handed over to the cloud computing service providers to complete. Not only saves the cost, but also does not need to expend the massive energy.The threat of network security is increasing, the network has a strong dependence on cloud computing is inevitable in the application process there are many security risks. In the traditional IT service solution, the vast majority of application software and data information is running or stored in the user's local physical equipment, in the user's absolute controllable range [2].

II. CP-ABE ALGORITHM AND SECRET SHARING SCHEME

A. CP-ABE Algorithm

Goyal et al.[3] proposed a new method of Attribute-Based Encryption (ABE) algorithm, which is a new method of access control under the condition of encryption. Later, ABE is divided into key policy attribute based encryption and CP-ABE. In the KP-ABE system, the access policy is made by DU. On the contrary, in the CP-ABE system, the access policy is formulated by DO, so the CP-ABE algorithm is more suitable for application in access control applications.

CP-ABE algorithm mainly includes the following four steps:

Step1 Generate a main (Key, MK) and public key PK.

Step2 C = Encrypt (PK, F, T): the data F using PK and access structure tree T encryption to get the cipher text C.

Step3 SK = Gen Key (MK, S): enter the main key MK and the property set S, the output of a private key (Key Secret, SK).

Step4 Decrypt (C, SK): as long as the SK contains the attribute set S to meet the access structure tree T, you can decrypt the encrypted C SK data obtained by F; otherwise into Step2.

B. Secret Sharing Scheme

Secret sharing scheme, also called (k, n) threshold scheme. As follows the secret s is divided into n blocks, S1,S2,…,Sn:Knowing any k or more Si data blocks, S can be easily calculated;Know any k - 1 or less Si data blocks, it is completely unable to get S.

This paper will use a method called All-Or-Nothing Transform-reed Solomon,AONT-RS. This method is to use All-Or-Nothing Transform,AONT to process before using Information Dispersal Algorithm to divide the data. The AONT scheme can be seen as a (n + 1, n + 1) threshold scheme, a document encoding divided into n + 1, can guarantee any number of slice less than the threshold will not able to decrypt data. IDA algorithm is a data slicing

algorithm, similar to the same SSS configuring a threshold, but the results of the slice will not increase with the factor. For example, the threshold scheme is (10, 15), then the total slice size is (15 /10) times as much as the original data[4].The related formulas are as follows,

$$f(x) = W_u \frac{r^3 \tan^3 \delta_0 \cos \delta_0}{\pi} x(\sin \delta_0 \cos \delta_0 + \frac{\pi}{2}) - \delta_0 (x2 + r^2 \tan^2 \delta_0)^2 \quad (1)$$

$$f(x) = W_0 \frac{3r^2}{\pi}(1 - \frac{x^2}{r^2}) \quad (2)$$

$$f(x) = W_0 \frac{2}{\sqrt{\pi^3} rx} e^{-4\frac{x^2}{r^2}}, n = 1 \quad (3)$$

$$f(x) = W_0 \frac{0.216x}{r} e^{-4\frac{x^6}{r^6}}, n = 3 \quad (4)$$

$$f(x) = W_0 \frac{2}{r^2} e^{-2\pi\frac{x^2}{r^2}}, n = 1 \quad (5)$$

$$f(x) = W_0 \frac{2}{r^2} e^{-2\pi\frac{x^2}{r^2}}, n = 2 \quad (6)$$

III. SECURE AND EFFICIENT CLOUD STORAGE SYSTEM FRAMEWORK

The secure and efficient cloud storage system is proposed in this paper, which is based on CP-ABE, and the AONT-RS scheme is used to optimize the performance. The framework of the cloud storage system is shown in Figure 1, where there are three participants: data owners, data users, and CSP. The main process of the system can be divided into three stages: data publishing, data retrieval and user revocation.

At the beginning of the data release phase, DO runs the Setup algorithm to generate public key PK and a master key MK. Then running the Key Gen algorithm to get the private key SK of each DU, and then SK send to DU through a secure channel. As the analysis in the introduction, the program needs to divide the original data F (assuming that the data F contains T words, each word has w bit) into pieces. Therefore DO will run the data slice (Splitting Data, DS) algorithm to segment the data. The corresponding algorithm is shown in algorithm 1. In this algorithm, the original data F first through the AONT method will code the original t into t + 1 words, of which the t + 1 word and CT+1 was used to check the integrity of the reconstructed data. The generated key K1 is used to reconstruct the phase, and is encrypted by the CP-ABE algorithm during the data release phase. And then through the (k, n) IDA algorithm will handle the data into n slices. Through Data Publishing, the data will be issued by DU to cloud random selection of N - (k - 1) a slice encryption, such unauthorized Du cannot recover the data through the rest of the (k - 1) slice, also in the cloud saving a copy of the data[4].

At present, although the cloud computing service providers through stable high-speed Internet connection allows users to access remote data storage, convenient and efficient access to services, but because of the cloud computing has virtual, large-scale, dynamic configuration and scalability and other characteristics, and has brought many security risks and challenges for the calculation of large-scale data storage service environment under the cloud. In order to improve the utilization rate of the storage efficiency and storage space, large data files are usually stored in the cloud computing service providers is split into a plurality of small blocks of data, location and storage of each data block of the user state is unknown, the user may doubt the integrity and consistency of your data file. As a key index to measure the data storage service, how to ensure the integrity and consistency of user data files stored in the geographic unknown huge server in the cluster, has always been a major problem in cloud computing and data storage services are facing. Especially after the Amazon Simple Storage Service and Google Docs service interruption and other accidents have occurred, users of cloud computing service providers is to save resources and reduce the cost of concealing safety accidents more had a great distrust. Users want to have a complete set of mechanisms so that they do not spend too much computing resources and time under the premise of a data file integrity and consistency of the ability to review. Related research has been carried out a long time ago, and has achieved good results in the design of efficiency, verifiability, query and recovery. Currently, there are two common solutions for data integrity and consistency: private audit and open audit. Private audit as the name implies is the users own commitment to the data file audit work, public audit is the audit trusted third party audit institutions to complete. Although the private audit because of its simple logic has higher efficiency of auditing, but public audit can not only provide safe and reliable data for the user, but also to a large extent for the user to save a lot of computing resources and time. In the cloud computing environment, users are unlikely to have a lot of time and energy to carry out frequent audit work on their own data files, will this time-consuming task by having reliable and complete audit protocol to solve the trusted third party audit plan to complete can be said to be a very good choice. Data storage research attention verification in terms of long ago already by the industry, the solution proposed by scholars in efficiency, verifiability, query and recoverability etc. also have achieved certain results. Unfortunately, most scholars' research is limited to the operation of static data files, many research results cannot meet the dynamic operation of data files frequently. The following will be scholars before the relevant research results are summarized. Based on the previous research results, this paper solves the verification problem of user data file integrity and consistency in cloud computing environment by using special tree structure of Merkle Hash Tree. Not only effectively support cloud computing basic data dynamic operation of all environments (including data add, delete and modify etc.), also give full consideration to the storage of data in a distributed environment the geographical location of influence on the

computational efficiency, the validation of the Merkle Hash based on Tree to make a certain contribution [5].

A. *Safety Assumption*

The scheme assumes that all communication channels do not exist in the case of large packet loss (communication channels include between USER and CDC, between CDC and TPA and between three parts of TPA and USER). At the same time, the TPA in the program is unbiased, fully trusted third party audit institutions, able to faithfully complete all tasks entrusted by USER. The security assumption of CDC in the scheme is slightly different from previous research, CDC is no longer completely non trust, but has a certain curiosity but can faithfully complete all tasks. CDC ensures that all parameters of their commitment to the correctness of calculation, no deceptive and non-repudiation, and unconditional response at any time for any user data file audit request. After all user data files are pretreated, there is no possibility of mutual interference between data blocks. In addition, the focus of the scheme is how to support the integrity and consistency verification of user data files stored in CDC and dynamic operation of data. Therefore, other securities issues such as user access control, data file recovery and so on are not within the scope of this chapter, this section will not be described in detail. Before the file is processed, USER will forward the data file storage request to CDC. CDC authentication of USER in accordance with its predefined access control rules. CDC authentication by legitimate users will get permission to file storage. In order to location only a block of data, for added position label to a size of 5 bytes for each data block (LTag), which is composed of a machine frame tag tag information sequence marking information, 2 bytes of the 1 byte and 2 byte node tag information composition. Sequence tag information record is the data blocks in all data blocks in the order number, frame marker information recording is the specific piece of data is stored in the data center frame number, node number specific server node tag information indicating the data block storage [6]. CDC maintains a LTag tag list for each data file, recording the LTag tag information for all data blocks of the file. File block sketch map shown in Fig.1.

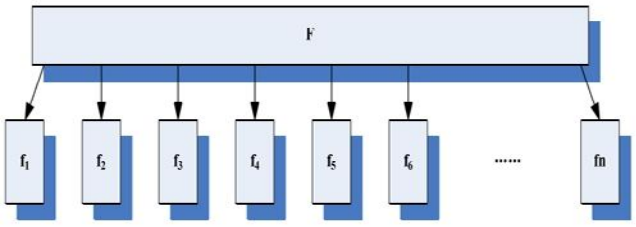

Figure 1. File block diagram

B. *Verification of Data Files*

In the process of file preprocessing, the Merkle Hash Tree of nodes, racks and files are constructed, and their corresponding root node values are calculated. All root node values will be the primary validation information for the solution. Program provides that the data file all validation information will not only be saved in CDC, but also in TPA backup. CDC is responsible for the latest authentication information to update the data file in real time through the communication network with TPA so that TPA can complete the data integrity and consistency verification of USER delegation. The program provides TPA file Merkle Hash Tree to the root node update data file with the latest LTag list information, stored data files of all the nodes Merkle Hash Tree root node value, the stored data files of all Merkle Hash Tree frame the root node value and the value of the stored data file. In addition, TPA clearly generates rules and methods for all validation information in a data file. First, CDC completes the USER data file storage, and submits to TPA the relevant verification data stored by USER. Data files before being deleted by USER, as long as the changes have occurred, CDC will be responsible for the data file to generate the latest validation information, and real-time updates with TPA. Then, CDC notification USER data file storage and verification information generation work has been completed, and told USER can begin to entrust TPA data file integrity and consistency of the verification. USER can choose immediately or try other time to communicate with the TPA, to entrust TPA to verify the integrity and consistency of data files. TPA from USER in the Audit Commission after the request, will carry out audit verification regularly or irregularly on the data file according to the requirement of USER, and all the audit operations retain verification log for USER days after the inspection. In the verification process, if the data file authentication fails, TPA will inform the USER by e-mail or SMS and other communications, while requiring CDC to recover data files and other remedies. Cloud computing environment, most of the data files carried out frequently three basic dynamic operations are: data insertion, data modification and data deletion. In this scheme, after the three dynamic operations of the data file, all the verification information related to the data file must be re generated by CDC.

IV. EXPERIMENTAL RESULTS

The encryption algorithm based on attribute by adding the user identity attribute description, use and gate, or gate and gate contains threshold function as constraint condition, significantly improve the ability of sharing data file, the system in a distributed environment, access control efficiency is better than the traditional use of unique identity label identity based encryption algorithm is very suitable. In the cloud computing environment data file sharing rate is very high. To research attention access control scheme in long before there has been the application of encryption algorithm based on attribute, the solution proposed by scholars in the permission revocation, threshold function support and proxy re encryption and other aspects also have achieved certain results. Unfortunately, most of the scholars of the research content are limited to the use of encryption algorithm based on attribute to solve the control problems of the traditional computing model in which access, most research are not able to meet the application requirements in cloud computing environment. The following first of

scholars before the relevant research results are summarized. The scheme assumes that all communication channels do not exist in the case of malicious packet loss . At the same time, the TPA in the program is unbiased, completely trusted third party audit institutions, able to faithfully complete all tasks entrusted by USER. The CDC in this paper is slightly different from the CDC in the previous scheme, although the CDC in the scheme has the curiosity, but can faithfully complete the task, no longer completely unreliable. All parameters of the CDC calculation, and bear the encryption key generation and distribution of work tasks, can guarantee the absolute correctness, no deceptive and non-repudiation. CDC can unconditionally respond to any USER file access request issued at any time, and strictly comply with the protocol formulated by the protocol key generation, change and distribution work. In addition, TPA can update real-time CDC access control and file information, CDC supervision. Cloud computing data storage security architecture, creative trusted third party audit agency into cloud computing access control operations. Therefore, the access control operation of the data file is slightly different from the traditional access control strategy. The program provides has write permissions for USER as long as the data file is modified, all have the data file access to the private key of the USER Merkle Hash Tree will file with the root node value change and failure, all USER access control permissions will be granted. When the USER key is proposed using failure data file access request, CDC will request to TPA for processing TPA, after confirming the identity of the USER, according to new data file access control permissions for the re issuance of private key operations. If USER still has access to the data files, then TPA will release a new private key for USER, so that USER can operate the corresponding data file; otherwise, TPA USER refused to file access request [10].

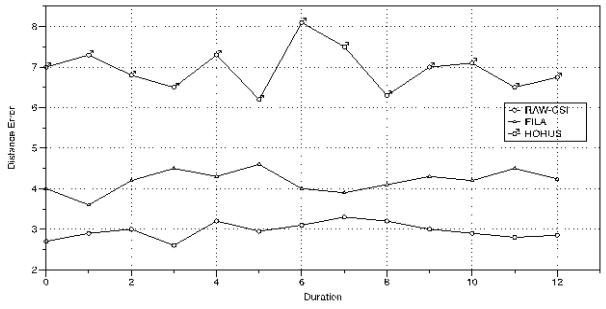

Figure 2. Mean Errors

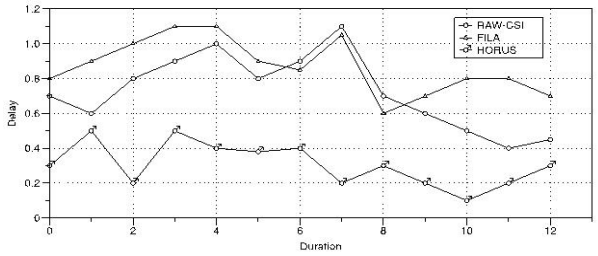

Figure 3. Mean Delays

Fig. 2 shows the comparison result among the three algorithms, RAW-CSI, HORUS and FILA. Apparently, in Fig. 2, RAW-CSI and FILA, both based on CSI, are superior to RSSI-as-fingerprint HROUS. Meanwhile, as shown in Fig.3, compared with FILA, because of the simplicity of RAW-CSI, its positioning delay is similar as FILA. i.e., RAW-CSI can reach a higher positioning accuracy than FILA without the breakdown of computing complication and positioning delay.

V. CONCLUSION

A secure and efficient cloud storage system is proposed in this paper. This system is an access control system based on CP-ABE, and puts forward a high efficient storage scheme based on data sharing and secret sharing, while only keeping a copy of the data. This scheme can significantly reduce the workload of DO and the storage space overhead of CSP, which can effectively promote the use of cryptography in the cloud storage system. At the same time, the security analysis proves that the system is safe. From the theoretical analysis and the actual test results, it can be seen that SECSS in the user revocation and storage space overhead is more efficient than OSCSS. Therefore, in the case where frequent and large amount of data is revoked, CSP and DO will benefit from that. Overall, the optimization scheme of this paper is to make an optimal balance between the system security and the overall overhead.

REFERENCES

[1] Kelejian H H and Prucha I R,“Estimation of simultaneous systems of spatially interrelated cross sectional equations,”Journal of Econometrics, vol.118, pp.27-50, Jul. 2014.

[2] Gebremariam G H, Gebremedhin T G and Schaeffer P V,“A simultaneous spatial panel data model of regional growth variation: An empirical analysis of employment, income, migration and local public services,”Computer Knowledge & Technology, vol.21, pp.34-38, Apr. 2015.

[3] Guangnan Zhang, Computer Technology.Beijing: Metallurgic Industry Press, 2016.

[4] Gebremariam G H, Gebremedhin T G, and Schaeffer P V,“Employment, income, migration and public services: A simultaneous spatial panel data model of regional growth,”Papers in Regional Science, vol.91, pp.275–297, Jun. 2014.

[5] Cherubini U, Mulinacci S and Gobbi F. Dynamic Copula Methods Infinance,New York:John Wiley & Sons, 2015.

[6] Fantazzini D. ,“Dynamic Copula Modelling for Value at Risk,” Frontiers in Finance & Economics, vol.45, pp.45-49, Jun. 2015.

[7] Patton A J. ,“Modeling Asymmetric Exchange Rate Dependence,” International economic review, vol. 34, pp.80 -85, Jul. 2014.

[8] Nelsen R B. An Introduction to Copulas.Springer,2014.

[9] Chu Wenkui and zhang Fengming,“Based on the military software security problems of the COTS study,”Journal of Systems Engineering and Electronics, vol. 3, pp.2166-2170, Aug. 2014.

[10] Tu Gang, zhang bo and Yang Fumin,“The research of embedded operating system transplant technology,”Computer application research, vol. 56, pp.83-85, Dec. 2014.