A Data Security Strategy for Cloud Computing

Xin Mingyuan [1,a], Wang Yong [2,b] and Fan lin [1,c]

[1.]College of Computer and Information Engineering Heihe University Heihe City, China
[2.]President's Office Heihe University Heihe City, China
[a.]xmy5686@163.com; [b.]8150767@qq.com; [c.]1135782701@qq.com

Keywords: Cloud computing; Data security; Policy

Abstract. This paper first analyzes the security problems of cloud computing background, in-depth analysis of the security challenges of cloud computing data security environment facing the development situation, summed up the domestic cloud data security technology and cloud data, put forward feasible security strategies, the safety verification in the simulated cloud environment, and puts forward the direction of later work.

Introduction

We are going through the era of deepening the reform of the Internet, where the network environment is constantly optimized and computer technology is developing rapidly. More and more data in work and life need to be processed with the help of network technology and computer technology, people's data storage methods from the earliest floppy disk, U disk to mobile hard disk to today's network disk and cloud storage, storage media capacity is multiplied, but the cost of storage is decreasing. Therefore, cloud storage is bound to become the key technology of Internet development and application. The concept of cloud enriches the available resources of the network and expands the network space, and injects new vitality into the application of network technology. With the rapid development of distributed network technology, virtual technology and computer technology, cloud computing technology has been produced. It has a new generation network commercial computer model which is compatible with multi-access point, resource sharing, extended dynamic and so on. Cloud computing solves the problems of uneven resource allocation, poor system computing power, insufficient storage capacity, and large investment in equipment and personnel management in traditional network computing mode. Cloud storage is the data base of cloud computing technology and the key technology to realize cloud computing applications. The rapid development of cloud technology has resulted in a new and efficient way of data management-service outsourcing. Users outsource data storage and resource computing services to third-party cloud servers, and can enjoy massive data storage and efficient data calculations at a very low cost, as well as offline access to data.

While enjoying the convenience and low cost of cloud computing for users in information calculation and storage, this new way of data calculation also presents a new challenge to data security. Most of the data that users upload to the cloud server contains key information such as the user's privacy data. If the user's personal information is not secure on the cloud, it will result in a large number of leaks of personal privacy information, resulting in a level of security incidents that will exceed the loss of any kind of powerful Trojan virus. Therefore, cloud security is the key to restrict the development of cloud computing.

Challenges to Cloud Data Security

Cloud data security technology runs through the various processes of cloud computing, combined with the general steps of information security, we divide cloud data security into four parts, namely, cloud data sharing security, cloud access control technology, searchable encryption technology and information integrity testing technology. Cloud computing security is different from the traditional way of computing, which makes the ownership and control of user data inconsistent, resulting in increased security risk of privacy information leakage. Cloud storage, as the core service provided

  DOI: 10.25236/icess.2019.227

by cloud computing, is a solution for cloud data sharing. The general practice of cloud data security is for users to encrypt data before uploading it to the server, and then to encrypt the data stored in the cloud server, usually the user of the data is not necessarily the owner of the data, then when the consumer uses the data, in order to protect the security of the data, A trusted agent decrypts the ciphertext and re-encrypts it and sends it to the consumer of the data to resolve the security issue of sharing the data. In addition, according to the user attribute feature encryption algorithm to distinguish between the data owner and the user, adapt to the dynamic strategy model of the Visitor dynamic transformation in the cloud environment to realize the identity difference access control technology. Data in the cloud server is encrypted data, the realization of efficient and accurate retrieval of encrypted data is the key to cloud computing, and in order to prevent the loss of user data due to system failure and other reasons, the necessary data integrity test scheme is needed.

Data sharing security issues.

The effective way to solve the leakage of privacy information is to encrypt the private information, in the traditional data calculation method, we generally choose symmetric or asymmetric encryption algorithm. And in the cloud computing data owners and users are inconsistent with the characteristics. Assuming that symmetric encryption algorithm is selected, no matter whether the key is generated on the cloud server or the client side, there will be users can not obtain control of the data and heavy key management problems. Assuming that the asymmetric encryption algorithm is chosen, the support of the public key infrastructure is required, and only the owner of the data shares the private key with the user in order to realize the fine-grained service of data calculation. Obviously this is a very dangerous practice. In the data sharing of cloud computing, the method of data re-encryption with third party agent and the attribute encryption algorithm based on the attributes of user identity attribute are usually adopted to solve this problem.Many experts at home and abroad have proposed concrete improvement algorithms for these two schemes. However, there are still defects, mainly in the following aspects: first, when the data owner uploads the ciphertext to the cloud Server Agent, the change of the user identity information of the cryptographic shared data and the cryptographic Data sharing policy update will cause the third party agent to have extremely complex data management problems and cause the risk of leaking the user's privacy data. Secondly, the attribute encryption algorithm relies too much on the trusted party to generate the key, which limits the environment of the cloud computing, but it can dynamically join the user to resist malicious conspiracy attacks, and it is easy to combine other security technologies, and the characteristics of reducing communication overhead are more suitable for the cloud computing environment.

Access control.

In terms of data access control, we propose the requirements of fine-grained and dynamic change permissions. In order to better protect the data owner's privacy information, access authorization to the data is refined to the file level, each access to each file with a different key encryption, authorization is the process of specifying the authorization of the key encrypted data. The dynamic changes of users in cloud computing, so that the authorization of files with the dynamic increase or decrease.The data access control algorithm in cloud computing has the CP-ABEalgorithm of KP-ABE and ciphertext strategy, which implies the decryption rule in the encryption algorithm, not only to please the PKI, but also to reduce the burden of key management in ciphertext access control. Therefore, ABE has been applied to the system implementation. However, because of the high cost and low efficiency of Abe encryption, when a ciphertext update operation occurs, it will produce a large number of operations that the data owner cannot afford. It is obviously not practical to assume that the client has control over the deployment of access control rules. Therefore, the dynamic management of multi-level access control is the place to be improved. Secondly, the access control in cloud computing is mostly focused on the control of reading cipher permissions, the lack of control of writing ciphertext data, and the failure to achieve comprehensive cloud storage security protection.

Searchable encryption technology.

In order to protect user privacy data from leakage, cloud server is mostly ciphertext data storage, only with the decryption key can read the data, decryption is done locally in the data consumer, so that a large amount of ciphertext information needs to communicate with the user between the third party, but, with the increase in the amount of ciphertext storage data, it is obviously difficult to achieve. So we tried to retrieve the ciphertext based on keywords on the cloud server side, returning only search results. However, this will lead to the leakage of sensitive information, so we explore a new cryptographic system, in the case of protecting user privacy information, in the cloud server side of the cryptographic information retrieval, return matching query results data, to achieve no decryption can be efficient and fast retrieval of ciphertext information services, And does not disclose the privacy information of the data acquisition party. Searchable encryption is divided into private key encryption (SSE) and public key encryption (peks) with keyword retrieval. In the old scheme, the data content of the data consumer is transparent to the server side, so that the cloud server can easily obtain the user's data information preference, and the third party cloud server is not a trusted party, so it is extremely dangerous for the untrusted party to obtain sensitive data time. Most of the schemes in public-key searchable cryptography are based on bilinear pairs, there are a lot of pairs of operations, encryption and decryption operations are too complex, and the retrieval efficiency needs to be improved. Private key searchable encryption technology, genetic key password System Key Management complexity of the problem, when applied to large data set retrieval, retrieval performance decline is obvious. The general solution is to calculate the index in advance, along with the security implications of the index.

Data Integrity Validation.

Data loss is the second most secure issue in cloud computing, according to a Cloud Security Alliance survey . Therefore, the integrity verification of cloud storage data is required before the user uses the data. The recyclability certificate (proofs of Retrievability) is referred to as Po R) and proof of data ownership (provable data possession, or PDP) can verify the integrity of cloud server data, but there are some limitations in these scenarios , for example, it is difficult to complete data validation when the amount of information that validates the data is large and the communication bandwidth is limited. At the same time, with the increase of verification computation, it brings unbearable storage overhead and complex key management to the verifier. And there is no precaution against the conspiracy and deception of cloud servers.

Data Security Policies under Cloud Computing

Model of the system s.

Based on the previous analysis, we found that the key to cloud computing security technology is the difficulty of resisting the conspiracy attacks of cloud servers and users, the private cloud as a trusted party, but limited by the use of domains, the public cloud is semi-trusted, so in the security model in this article we divide the cloud into trusted private and semi-trusted public clouds, Together with the owner of the data and the user of the data share. Data store will need cloud storage service data encryption after uploading to the cloud server, ciphertext data c={c1,c2 ... Cn}, in order to adapt to the characteristics of dynamic change of cloud users, to achieve dynamic sharing security control based on attribute base and user's fine-grained access control, data owner and private cloud between the use of symmetric cryptosystem, private cloud and public cloud between the key set random selection key re-encryption and then send ciphertext to the public cloud, The fuzzy matching attribute base cryptosystem is adopted between the public cloud and the user, and in order to realize the cryptographic-based retrieval function, the data owner generates the fuzzy matching set of keywords according to the data, which is stored in the private cloud. For the user of the data, the fuzzy set of the keyword is calculated after the Retrieval keyword information is transmitted to the private cloud. The private cloud then sends the keyword Fuzzy match set to the public cloud, and the public cloud finally uses the index to complete the retrieval service of the user data, and returns the ciphertext retrieval results to the private cloud co-user download.

Privacy protection scheme of agent re-encryption based on attribute base.

The scheme is based on the idea of agent re-encryption technology, re-encrypts the cipher encrypted by the data owner into the secret of private cloud encryption, and the data has been in ciphertext state throughout the transformation process. The agent re-encryption process uses the private key of the data owner and generates the attribute based public key algorithm to ensure the privacy of the cloud data. According to the design idea, the algorithm can be divided into the following process. Initialization Set u to a shared cloud property set, p as prime number, and ZP to attribute related elements. The TI on the ZP is randomly selected as the private key, and the public key is:

$$\left\{T_1 = g^{t1}, \cdots, T_{|U|} = g^{t|u|}\right\} \tag{1}$$

The build agent re-encrypts the private key. For a subset on U‘ U， Generation satisfies q(0)=y of q(x)。

$$D_i = g^{\frac{q(i)}{ti}} \tag{2}$$

Encryption. The data owner encrypts with the private key to form the initial encryption, which is then re-encrypted by the public cloud using TI. The G2 is a bilinear pairing operation, and S is a random value on the G2 of Set U ' and plaintext M. The ciphertext is:

$$E = (U', \mathrm{E}' = \mathrm{MY}^{\mathrm{S}}, \{E_i = T_i^s\}_{i \in u'}) \tag{3}$$

Decrypt. Data consumers download ciphertext and decrypt it locally, often using decryption keys and re-encryption as different cryptographic algorithms.

A Searchable Scheme of Fuzzy Keyword Based on Index

The idea of the scenario is that all and consumers of the data send the retrieved exact keyword to the private cloud, the index number corresponding to the fuzzy set that calculates the keyword on the private cloud, and loads it to the client after using the index number in the public cloud to retrieve the ciphertext results. The specific algorithm is described as:

Index building.

The data owner is responsible for indexing the generated ciphertext and keyword ambiguity sets and uploading them to the private cloud. For each keyword, the data owner uses the key SD to compute all trap values.

Apply for a search.

To search for a document that contains a keyword m, the data consumer first sets the search scope D and generates the Blur set Sm， d; and then to each element of the blur set Calculate Trap Gate Values.

Document search.

The user sends the Search keyword request to the private cloud, and the fuzzy match to the search index list is sent to the secret text of the shared cloud to retrieve the query results, and the ciphertext is returned to the user requesting the search service for download.

Summary

In this paper, the proposed scheme is a performance optimization scheme based on the agent re-encryption algorithm, so the privacy of the scheme is secure, in the ciphertext searchable scheme, the main security problem is index disclosure problem, but the index generation in the scheme is on the private cloud trusted end, so there is no risk of privacy data leakage. However, this scheme increases the process and computation of ciphertext processing, and in the case of massive data, there is no

validation experimental data for computational efficiency. Next, we will simulate the massive data processing environment to evaluate the efficiency and anti-attack capability of the scheme..

Acknowledgment

Higher education teaching reform research project SJGY0213 in 2017; basic research business fee research project 2017-KYYWF-0353 in Heilongjiang Education Department; Backbone teachers plan in Heihe University; Heilongjiang Province youth talent plan UNPYSCT-2017104 in 2017.

References

[1] Wei L, Reiter MK. Ensuring file authenticity in private DFA evaluation on encrypted files in the cloud. In: Crampton J, Jajodia S, Mayes K, eds. Proc. of the 18th European Symp. on Research in Computer Security (ESORICS 2013). Heidelberg: Springer-Verlag, 2013. 147163. [doi: 10.1007/978-3-642-40203-6_9]

[2] Pattuk E, Kantarcioglu M, Lin ZQ, Ulusoy H. Preventing cryptographic key leakage in cloud virtual machines. In: Proc. of the 23rd USENIX Security Symp. (SEC 2014). Berkeley: USENIX Association, 2014. 703-718.

[3] Varadarajan V, Ristenpart T, Swift M. Scheduler-Based defenses against cross-VM side-channels. In: Proc. of the 23rd USENIX Security Symp. (SEC 2014). Berkeley: USENIX Association, 2014. 687-702.

[4] Rocha F, Correia M. Lucy in the sky without diamonds: Stealing confidential data in the cloud. In: Proc. of the 41st IEEE/IFIP Int'l Conf. on Dependable Systems and Networks Workshops (DSNW 2011). Washington: IEEE Computer Society, 2011. 129-134. [doi: 10.1109/DSNW.2011.5958798]

[5] Cloud Security Alliance. The notorious nine cloud computing top threats in 2013. TechnicalReport,2013.http://www.chinacloud.cn/upload/2013-03/13030711513081.pdf