Survey of Privacy Data Protection Technology in Cloud Computing

Yang Wang
Hangzhou Dianzi University, China

ABSTRACT: Cloud computing is very populate because of computing and storage resources. Users are likely to upload their privacy data to the Cloud computing platform, which may easily suffer security threats and vulnerabilities, including secure data transmission, data loss, data breach, malicious insiders, account/service hijacking. In order to protect privacy data, kinds of technologies are proposed including data encryption, data isolation, access control, cipher text retrieval, data auditing and copyright protection.

1 INSTRUCTIONS

Cloud computing platform has a large of computing and storage resources, and accumulates a large amount of privacy data from kinds of users. They store and share important data thorough the Cloud platform. So the data threat is more complex than in the traditional computing system. In 2009, Gartner, an international consulting company, found that multi users shared service mode usually brought new challenges to data security, and the Cloud service providers did not fully guarantee the security of the data.

In recent years, more and more Cloud computing security events are exposed. In March 2011, massive user data was lost with Gmail mailbox, where the content and chat records in mails were deleted. In 2013, Amazon's Cloud computing services stopped many times so that lots of social applications had been affected. In April 2013, Apple's iCloud suffered a broken network, as a result, GameCenter and itunes could not provide services. More seriously, in June 2013, "prism" event is exposed, which showed that large of privacy data was likely to be leaked or mining analysis from Cloud storage center, such as mail, instant messages, videos, photos and other. Therefore, the privacy data protection technology of Cloud computing has aroused attention from many researchers.

2 DEFINITION AND APPLICATION OF CLOUD COMPUTING

Cloud computing is a new computing model, which provides users with rich computing and storage resources in the form of service [1]. National Institute of standards and Technology Research Institute (NIST) defines the Cloud computing that Cloud computing is a kind of Internet-based computing, which provides shared processing resources to computers and other devices on demand. It is also a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources, such as networks, servers, storage, applications and services [2]. Cloud computing can save the cost of the users, and improve the utilization of computing resources. Therefore, Cloud computing is widely used and develops fast.

Cloud computing has five remarkable characteristic: on-demand service, broadband Internet access, virtual resource pool, fast and flexible architecture and measurement services. While Cloud computing has three service models: infrastructure as a service, platform as a service and software as a service. In addition, according to the nature of services and business model, Cloud computing is divided into three categories: private cloud, public cloud and hybrid cloud [3][4].
3 DATA SECURITY ISSUES IN CLOUD COMPUTING

In 2016, the Cloud Security Alliance (CSA) announces the top Cloud computing threats and vulnerabilities in an enterprise environment, which consider 12 cloud vulnerabilities, including secure data transmission, data loss, data breach, malicious insiders, account/service hijacking [5]. Based on the characteristics of the Cloud computing and security problems, we summarize the main security threats and the corresponding solutions in the Cloud computing platform.

3.1 Cloud service providers may steal and change the privacy data stored in Clouding computing platform

The transparency of Cloud computing separates the data from the platform infrastructure, and shield the underlying details to the common users. The service model allows users to upload data to the cloud server, which makes the cloud service providers to access user data easily. At worst, the untrusted cloud service providers even steal and change the data without the permission of users. In addition, malicious insiders or system failures often do not follow the best security guidelines and do not implement the security policy, which will lead to the risk of data manipulation.

3.2 Unauthorized users may access the privacy data stored in Clouding computing platform

A large number of privacy data are concentrated in the cloud server. If the Cloud computing platform does not have a reasonable authentication mechanism, it would be possible for malicious unauthorized users to fake legitimate users then to conduct a variety of illegal operations, such as stealing the privacy data, modifying important configuration information, and opening the user's privacy information. They bring huge losses to legitimate users. Therefore, identity authentication is an important prerequisite for users to access the data in the Cloud computing platform. Cloud service providers should provide strict authentication mechanism.

3.3 Cloud service providers may data mine the privacy in the process of data using

Pearson divided the privacy information into three kinds: personal identification information, personal sensitive information and personal data information [6]. Personal identification information contains name and address. Personal sensitive information contains job and religion. Personal data information contains photo and files. In some typical Cloud computing services, such as group purchase, electronic commerce website, cloud search, may get identity information, usage habits of users based on mining and analysis of the data. Therefore, Cloud computing platform should take measures to prevent cloud service providers to collect user's sensitive information.

4 DATA PRIVACY PROTECTION TECHNOLOGY IN CLOUD COMPUTING

The security of data in the Cloud computing platform is related to data confidentiality, integrity and availability, as well as user authentication, authorization access and modification, etc. At present, the data security technologies of Cloud computing platform mainly include data encryption, data isolation, access control, cipher text retrieval, data auditing and copyright protection [7].

4.1 Data encryption

Data encryption is a most effective way to protect privacy data security. According to the use frequency, data in the Cloud computing platform is divided into two kinds: static data and dynamic data [8]. Accordingly, data encryption mechanism also has two forms: static encryption and encryption. The encryption scheme in modern cryptography includes symmetric encryption and asymmetric encryption, and the former is further classified as block cipher, stream cipher and hash function. Symmetric encryption algorithms have the same encryption keys and decryption keys, such as DES, AES, IDEA, RC4, RC5, RC6. Asymmetric encryption algorithms have different private key and public key, such as RSA, EIGamal algorithm based on the discrete logarithm.

For the key management, it is generally adopted hierarchical key management. It only needs to keep a small number of keys in data nodes. The efficiency is higher.

4.2 Access control

Data protection can through access control list (ACL) to realize the control of user access to encrypted data. But this method is lack of flexibility and extension. It is proportional between the number of users and the overhead of data encryption, which is not suitable in the Cloud computing platform. In order to achieve more flexible access to encrypted data, Waters and Sahai, based on the identity based encryption, proposed Based Encryption Attribute (attribute based encryption, ABE) mechanism [9]. In ABE mechanism, the access structure is introduced in the idea of public key encryption. The key or cipher text is generated according to the access
structure, so that the user can decrypt the data if they meet the specified conditions.
With deepening of ABE research, the mainly researches focus on key escrow, decryption efficiency, and attribute revocation.

4.3 Cipher text retrieval

Cipher text retrieval is based on the method of index or the cipher text. In order to improve the efficiency of retrieval and protect the privacy of users, the encrypted [10] (Encryption Searchable) can be searched. At present, according to the different construct algorithms, Cipher text retrieval mechanism is divided into two categories: Searchable Symmetric Encryption (SSE) mechanism, which is based on symmetric cryptography algorithms, and Public Key Encryption with Keyword Search (PEKS), which is based on public key cryptography algorithm.

In order to support the multi user encrypted retrieval in Cloud computing platform, and support the search conditions of complex logic, the Based Searchable Encryption [11] (ABSE) mechanism is proposed. ABSE mechanism defines keyword with attributes, which can express the search conditions better. When users search the keywords, search evidence is generated according to the attributes and then is sent to the cloud computing platform. The Cloud computing platform immediately finds the correct results through all the cipher texts.

4.4 Data transformation

Even if users do not directly release their private data to the Cloud platform, cloud service providers may still get user's privacy information by data mining. In order to avoid this situation, privacy protection technology based on data transformation and anonymous can be used. In the field of data anonymous, privacy, K- anonymity model and t-closeness model [12]. Those models can transform initial data to implicit content, which cannot be readable. After data transformation, privacy information cannot be identified easily. And then the privacy data is protected to be leakage.

5 CONCLUSION

The Cloud computing is developed rapidly. It not only allows users to upload their private data to the Cloud storage platform, but also let users visit the data at any time and place. This method brings serious data security and privacy disclosure issues. Therefore, it is important to protect data confidentiality and integrity in the Cloud computing platform. On the other hand, data security is also a prerequisite for users to store and share data in the Cloud computing platform. Although there have been some theory results about private data security, it is still in the initial stage. We will work on the data protection in the Cloud computing in the future.

REFERENCES