Attribute Based Storage Supporting Secure Deduplicationof Encrypted Data In Cloud

PULAVARTHI S V SAIRAM GUPTHA,

MR. NAGA SRINIVASA RAO
[1]**PG STUDENT, DEPT OF MCA**

[2] **Asst. Prof**, Dept of MCA

SREE KONASEEMA BHANOJI RAMARS P.G. COLLEGE (AMALAPURAM, Andhra Pradesh)

ABSTRACT

Attribute-based encryption (ABE) has been widely used in cloud computing where a data provider outsources his/her encrypted data to a cloud service provider, and can share the data with users possessing specific credentials (or attributes). However, the standard ABE system does not support secure deduplication, which is crucial for eliminating duplicate copies of identical data in order to save storage space and network bandwidth. In this paper, we present an attribute-based storage system with secure deduplication in a hybrid cloud setting, where a private cloud is responsible for duplicate detection and a public cloud manages the storage. Compared with the prior data deduplication systems, our system has two advantages. Firstly, it can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion. In addition, we put forth a methodology to modify a cipher text over one access policy into cipher texts of the same plaintext but under other access policies without revealing the underlying plaintext.

I. INTRODUCTION

Cloud computing greatly facilitates data providers who want to outsource their data to the cloud without disclosing their sensitive data to external parties and would like users with certain credentials tobe able to access the data . This requires data to be stored in encrypted forms with access control policiessuch that no one except users with attributes (or credentials) of specific forms can decrypt the encrypted data. An encryption technique that meets this requirement is called attribute-based encryption(ABE) , where a user's private key is associated with an attribute set, a message is encryptedunder an access policy(or access structure) over a set of attributes, and a user can decrypt a ciphertext with his/her private key if his/her set of attributes satisfies the access policy associated with this ciphertext. However, the standard ABE system fails to achieve secure deduplication , which is a technique to save storage space and network bandwidth by eliminating redundant copies of the encrypted data stored in the cloud. On the other hand, to the best of our knowledge, existing constructions for secure deduplication are notbuilt on attribute-based encryption. Nevertheless, since ABE and secure deduplication have been widelyapplied in cloud computing, it would be desirable to design a cloud storage system possessing bothproperties.

We consider the following scenario in the design of an attribute-based storage system supporting secure deduplication of encrypted data in the cloud, in which the cloud will not store a file more than once even though it may receive multiple copies of the same file encrypted under different access policies. A data provider, Bob, intends to upload a file M to the cloud, and share M with users having certain credentials. In order to do so, Bob encrypts M under an access policy A over a set of attributes, and uploads the corresponding cipher text to the cloud, such that only users whose sets of attributes satisfying the access policy can decrypt the cipher text. Later, another data provider, Alice, uploads a cipher text forthe same underlying file M but ascribed to a different access policy A0. Since the file is uploaded in an encrypted form, the cloud is not able to discern that the plaintext corresponding to Alice's cipher text is the same as that corresponding to Bob's, and will store M twice. Obviously, such duplicated storage wastes storage space and communication bandwidth.

II. LITERATURE SURVEY

Cloud computing, a convenient way of accessing services, resources and applications over the Internet, shifts the focus of industries and organizations away from the deployment and day to-day running of their IT facilities by providing an on-demand, self-service, and pay-as-you-go business model. It is, therefore, unsurprising that cloud computing has continued to increase in popularity in recent times. While cloud computing provides various benefits to users, there are underlying security and privacy risks. For example, multi-tenancy, resource pooling and shareability features can be exploited by cybercriminals and any one with a malicious intent, to the detriment of both cloud users and cloud service providers. It is unsurprising, then, that cloud computing has emerged as a salient area of inquiry for security researchers. For example, when user data (e.g. documents, videos and photos) are uploaded or stored in a cloud computing service, the data owners are unlikely to know the path of the transmitted data or whether the data are being collected and analyzed by a third party, including a government agency (see the revelations by Edward Snowden [8]). As posited by Choo and Sarre [6], it is important to strike a balance between privacy, legitimate surveillance and lawful data access, in order to ensure that the privacy of innocent individuals will not be compromised (e.g. that fine-grained aspects of an individual’s life cannot be derived or inferred from the intelligence collection and analysis). A particularly promising approach to achieve security and privacy in this new computing paradigm is through cryptography [19,21]. For example, as noted by Yang et al. [25], to ensure the security and privacy of user data, specifically against an untrusted cloud service provider, one could encrypt the data prior to uploading and storing the data in the cloud. This special issue is dedicated to providing both scientists and practitioners with a forum to present their recent research on the use of novel cryptographic techniques to improve the security and privacy of the underlying cloud architecture or ecosystem, particularly research that integrates both theory and practice. For example, how do we design an efficient cloud cryptography system that offers enhanced security and/or privacy without compromising on usability and performance? In the sequel, we briefly survey the content of papers in this special issue.

Cloud forensics: State-of-the-art and future directions

Cloud log forensics (CLF) mitigates the investigation process by identifying the malicious behaviour of attackers through profound cloud log analysis. However, the accessibility attributes of cloud logs obstruct accomplishment of the goal to investigate cloud logs for various susceptibilities. Accessibility involves the issues of cloud log access, selection of proper cloud log file, cloud log data integrity, and trustworthiness of cloud logs. Therefore, forensic investigators of cloud log files are dependent on cloud service providers (CSPs) to get access of different cloud logs. Accessing cloud logs from outside the cloud without depending on the CSP is a challenging research area, whereas the increase in cloud attacks has increased the need for CLF to investigate the malicious activities of attackers. This paper reviews the state of the art of CLF and highlights different challenges and issues involved in investigating cloud log data. The logging mode, the importance of CLF, and cloud log-as-a-service are introduced. Moreover, case studies related to CLF are explained to highlight the practical implementation of cloud log investigation for analysing malicious behaviour. The CLF security requirements, vulnerability points, and challenges are identified to tolerate different cloud log susceptibilities.

We identify and introduce challenges and future directions to highlight open research areas of CLF for motivating investigators, academicians, and researchers to investigate them.

Cloud based data sharing with fine-grained proxy re-encryption

Conditional proxy re-encryption (CPRE) enables fine-grained delegation of decryption rights, and has many real-world applications. In this paper, we present a ciphertext-policy attribute based CPRE scheme, together with a formalization of the primitive and its security analysis. We demonstrate the utility of the scheme in a cloud deployment, which achieves fine-grained data sharing. This application implements cloud server-enabled user revocation, offering an alternative yet more efficient solution to the user revocation problem in the context of fine-grained encryption of cloud data. High user-side efficiency is another prominent feature of the application, which makes it possible for users to use resource constrained devices, e.g., mobile phones, to access cloud data. Our evaluations show promising results on the performance of the proposed scheme.

III. SYSTEM ANALYSIS EXISTING SYSTEM

When a user uploads data that already exist in the cloud storage, the user should be deterred from accessing the data that were stored before he obtained the ownership by uploading it (backward secrecy)2. These dynamic ownership changes may occur very frequently in a practical cloud system, and thus, it should be properly managed in order to avoid the security degradation of the cloud service. In the former approach, most of the existing schemes have been proposed in order to perform a PoW process in an efficient and robust manner, since the hash of the file, which is treated as a “proof” for the entire file, is vulnerable to being leaked to outside adversaries because of its relatively small size. a data owner uploads data that do not already exist in the cloud storage, he is called an initial uploader; if the data already exist, called a subsequent uploader since this implies that other owners may have uploaded the samedata previously, he is called a subsequent uploader.

DISADVANTAGES:-

User deduplication on the client-side, cannot generate a new tag when they update the file. In this situation, the dynamic Ownerships would fail. As a summary, existing dynamic Ownerships cannot be extended to the multi-user environment. Whenever data is transformed, concerns arise about potential loss of data. By definition, data deduplication systems storedata differently from how it was written. As a result, users are concerned with the integrity of their data. One method for deduplicating data relies on the use of cryptographic hash functions to identify duplicate segments of data. If two different pieces of information generate the same hash value, this is known as a collision. The probability of a collision depends upon the hash function used, and although theprobabilities are small, they are always non zero.

Proposed System

This Project the goal of saving storage space for cloud storage services also is used for secure deduplication .but several process have been this

same concept for deduplication. however this project flow some different modules in there . In this case, if two users upload the same file, the cloud server can discern the equal ciphertexts and store. only one copyof them. This process some authentication available insome issue for security purpose . through this process for ensure secured deduplication. A owner wants to outsource data to the cloud and share it with users possessing certain credentials. The Attribute Authority issues every user a decryption key associated with users set of attributes. which is considered to be the most important challenge for efficient and secure cloud storage services in the environment where ownership changes dynamically. Every time data provider upload file checking from cloud for save storage purpose . Most of the schemes have been proposed to provide data encryption, whilestill benefiting from a deduplication technique. every user get secured key form admin for security purpose

.user can not take any key he can not download chipertext file .they can download only encrypted data.every details manage and maintain by Attribute authority. In this way, any user who downloads the file, after decryption, can check the correctness of the decrypted plaintext by matching it to the corresponding tag. To keep the notation succinct, we use c to denote the combination of the encrypted data and the corresponding access structure

ADVANTAGES:-

The system has two advantages. Firstly, it can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. Secondly, it achieves the standard notion of semantic security for data confidentiality while existing systems only achieve it by defining a weaker security notion. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but underother access policies without revealing the underlyingplaintext.

I. IMPLEMENTATIONModules Description

These are the following modules

- DATA OWNER
- CLOUD SERVER
- AUTHORITY
- END USER

Data Owner

In this module, initially the data owner has toregister to the cloud server and get authorized. After the authorization from cloud data owner will encryptand add file to the cloud server where in after the addition of file data owner requests the content key and the master secret key to the authority for the file he uploaded and finds Find deduplication ,only after the keys generated the file is uploaded to the cloud server. After the uploading of the file the data owner will have to provide download and the search permission for individual file for the users to performsearch and download.

Cloud Server

The cloud server manages a cloud to providedata storage service. Data owners encrypt their data files and store them in the cloud for sharing with cloudEnd users. To access the shared data files users will request the permission of content key and the MSK master secret key. And the cloud will provide the permission .and also views all the transactions and attackers related to the files.

Authority

Authority generates the content key and thesecret key requested by the end user.
Authority can view all files with the content key and master secret key generated with the corresponding data owner details of the particular file.

End User

User has to register and login for accessing the files in the cloud. User is authorized by the cloud to verify the registration. User has to request for the MSK master secret key and content key to download the file. User can only download and search the file ifthe data owner of the particular file has provided the permissions.

II. CONCLUSION

Attribute-based encryption (ABE) has been widely used in cloud computing where data providers outsource their encrypted data to the cloud and can share the data with users possessing specified credentials. On the other hand, deduplication is an important technique to save the storage space andnetwork bandwidth, which eliminates duplicate copies.

of identical data. However, the standard ABE systemsdo not support secure deduplication, which makesthem costly to be applied in some commercial storage services. In this paper, we presented a novel approach to realize an attribute-based storage system supporting secure deduplication. Our storage system is built under a hybrid cloud architecture, where a private cloud manipulates the computation and a public cloud manages the storage. The private cloud is provided with a trapdoor key associated with the correspondingcipher text, with which it can transfer the cipher text over one access policy into cipher texts of the same plaintext under any other access

policies without beingaware of the underlying plaintext. After receiving a storage request, the private cloud first checks t he validity o f the uploaded item through the attached proof. If the proof is valid, the private cloud runs a tagmatching algorithm to see whether the same data underlying the cipher text has been stored. If so, whenever it is necessary, it regenerates the cipher textinto a cipher text of the same plaintext over an access policy which is the union set of both access policies. The proposed storage system enjoys two major advantages. Firstly, it can be used to confidentially share data with other users by specifying an access policy rather than sharing the decryption key. Secondly, it achieves the standard notion of semantic security while existing deduplication schemes only achieve it under a weaker security notion.

REFERENCES

[1] Hui Cui, Robert H. Deng, Yingjiu Li, and Guowei Wu, “Attribute-Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud”.

[2] D. Quick, B. Martini, and K. R. Choo, CloudStorageForensics. Syngress Publishing/Elsevier,2014.[Online].Available:http://w ww.elsevier.com/books/cloud-storageforensics/quick/978-0-12-419970-5

[3] K. R. Choo, J. Domingo-Ferrer, and L. Zhang, “Cloud cryptography: Theory, practice and future research directions,” Future Generation Comp. Syst., vol. 62, pp. 51–53, 2016.

[4] K. R. Choo, M. Herman, M. Iorga, and B. Martini,“Cloud forensics: State-of-the-art and future.

directions,” Digital Investigation, vol. 18, pp. 77–78,2016.

[5] Y. Yang, H. Zhu, H. Lu, J.Weng, Y. Zhang, and K.
R. Choo, “Cloud based data sharing with fine-grainedproxy re-encryption,” Pervasive and Mobile Computing, vol. 28, pp. 122–134, 2016.

[6] D. Quick and K. R. Choo, “Google drive: Forensicanalysis of data remnants,” J. Network and Computer Applications, vol. 40, pp. 179– 193, 2014.

[7] A. Sahai and B. Waters, “Fuzzy identity-based encryption,” in Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings, ser. Lecture Notes in Computer Science, vol. 3494. Springer, 2005, pp. 457–473.

[8] B. Zhu, K. Li, and R. H. Patterson, “Avoiding the disk bottleneck in the data domain deduplication file system,” in 6th USENIX Conference on File and Storage Technologies, FAST 2008, February 26- 29, 2008, San Jose, CA, USA. USENIX, 2008, pp. 269– 282.

[9] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secure deduplication,” in Advances in Cryptology - EUROCRYPT 2013, 32nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Athens, Greece, May 26- 30, 2013. Proceedings, ser. Lecture Notes in ComputerScience, vol. 7881. Springer, 2013, pp. 296–312.

[10] M. Abadi, D. Boneh, I. Mironov, A. Raghunathan, and G. Segev, “Message-locked encryption for lock-dependent messages,” in Advances in Cryptology - CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I, ser. Lecture Notes in Computer Science, vol. 8042. Springer, 2013, pp. 374–391.

[11] S. Keelveedhi, M. Bellare, and T. Ristenpart, “Dupless: Serveraided encryption for deduplicated storage,” in Proceedings of the 22th USENIX SecuritySymposium, Washington, DC, USA, August 14-16, 2013. USENIX Association, 2013, pp. 179–194.

[12] M. Bellare and S. Keelveedhi, “Interactive message-locked encryption and secure deduplication,” in Public-Key Cryptography – PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 – April 1, 2015, Proceedings, ser. Lecture Notes in

Computer Science, vol. 9020. Springer, 2015,pp. 516–538.