Attribute Based Storage Supporting Secure Deduplication of Encrypted Data in Cloud

[1] M.SURESH KUMAR NAIDU
M.Sc. (Computer Science)
Besant Theosophical College, Madanapalle.

[2] Dr.P.VEERAMUTHU
Assistant Professor
Besant Theosophical College, Madanapalle.

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 inorder 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 ciphertext over one access policy into ciphertexts of the same plaintext but under other access policies

without revealing the underlying plaintext.

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 same data 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 store data 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 the probabilities are small, they are always non zero.

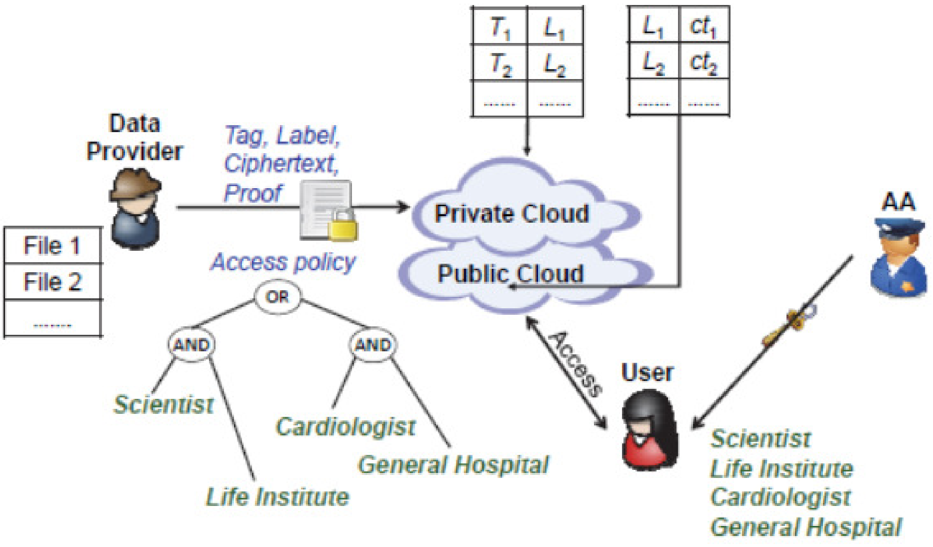
PROPOSED SYSTEM:

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 copy of them. This process some authentication available in some 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, while still benefiting from a deduplication technique. every user get secured key form admin for security purpose .user can not take any key he can not downloadchipertext 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:-

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 under other access policies without revealing the underlying plaintext.

System architecture of attribute-based storage with secure Deduplication.:



Modules:

In this project we have following Four modules .

i).Data Provider

ii).Cloud

iii).Deduplicaion

iv).Attribute Authority

Data Provider:-

Data provider uploading file to cloud with tag , label and security key , the proposed scheme guarantees data integrity against any tag inconsistency attack. Thus, security is enhanced in the proposed scheme.

CLOUD STORAGE:-

Secure Deduplicationwith the goal of saving storage spacefor cloud storage services, Douceur et al the first solution for balancing confidentiality and efficiency in performing deduplication called convergent encryption, where a message is encrypted under a message-derived key so that identical plaintexts are encrypted to the same ciphertexts. In this case, if two users upload the same file, the cloud server can discern the equal ciphertexts and store only one copy of them. which may violate the privacy of the data if the cloud server cannot be fully trusted . This is a client who owns data, and wishes to upload it into the cloud storage to save costs. A data owner encrypts the data and outsources it to the cloud storage with its index information, that is, a tag.

Deduplication:-

Data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. Related and somewhat synonymous terms are intelligent (data) compression and single-instance (data) storage. This technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. In the deduplication process, unique chunks of data, or byte patterns, are identified and stored during a process of analysis.Deduplication techniques take advantage of data similarity to identify the same data and reduce the storage space. In contrast, encryption algorithms randomize the encrypted files in order to make ciphertext indistinguishable from theoretically random data.

Attribute Authority:

The AA issues every user a decryption keyassociated with user set of attributes At the user side, each user can download an item, and decrypt the ciphertext with the attribute-based private key generated by the AA if this user's attribute set satisfies the access structure.

ALGORITHMS:-

RSA Algorithm:

RSA is an algorithm used by modern computers to encrypt and decrypt messages. It is an asymmetric cryptographic algorithm. Asymmetric means that there are two different keys. This is also called public key cryptography, because one of them can be given to everyone. The other key must be kept private. It is based on the fact that finding the factors of an integer is hard (the factoring problem). RSA stands for Ron Rivest, Adi Shamir and Leonard Adleman, who first publicly described it in 1978. A user of RSA creates and then publishes the product of two large prime numbers, along with an auxiliary value, as their public key. The prime factors must be kept secret. Anyone can use the public key to encrypt a message, but with currently published methods, if the public key is large enough, only someone with knowledge of the prime factors can feasibly decode the message.

I).ENCRYPTION ALGORITM

II).DECRYPTION ALGORITHM

III).DEDUPLICATION

ENCRYPTION ALGORITM:-

Encryption allows information to be hidden so that it cannot be read without special knowledge (such as a password). This is done with a secret code or cypher. The hidden information is said to be encrypted.

DECRYPTION ALGORITHM:-

Decryption is a way to change encrypted information back into plaintext. This is the decrypted form. The study of encryption is called cryptography. Cryptanalysis can be done by hand if the cypher is simple. Complex cyphers need a computer to search for possible keys. Decryption is a field of computer science and mathematics that looks at how difficult it is to break a cyphe

CONCULUTION:

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 inorder 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. That can be used to confidentially share data with users by specifying access policies rather than sharing decryption keys. In addition, we put forth a methodology to modify a ciphertext over one access policy into ciphertexts of the same plaintext but under other access policies without revealing the underlying plaintext.

REFERENCES:

[1] J. Crowcroft, “On the duality of resilience and privacy,” in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 471, no. 2175. The Royal Society, 2015, p. 20140862.

[2] A. Bessani, M. Correia, B. Quaresma, F. Andre, and P. Sousa, “Depsky: ´ dependable and secure storage in a cloud-of-clouds,” ACM Transactions on Storage (TOS), vol. 9, no. 4, p. 12, 2013.

[3] H. Chen, Y. Hu, P. Lee, and Y. Tang, “Nccloud: A network-coding-based storage system in a cloud-of-clouds,” 2013.

[4] T. G. Papaioannou, N. Bonvin, and K. Aberer, “Scalia: an adaptive scheme for efficient multi-cloud storage,” in Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis. IEEE Computer Society Press, 2012, p. 20.

[5] Z. Wu, M. Butkiewicz, D. Perkins, E. Katz-Bassett, and H. V. Madhyastha, “Spanstore: Cost-effective geo-replicated storage spanning multiple cloud services,” in Proceedings of the Twenty-Fourth ACM Symposium on Operating Systems Principles. ACM, 2013, pp. 292–308.

[6] G. Greenwald and E. MacAskill, “Nsa prism program taps in to user data of apple, google and others,” The Guardian, vol. 7, no. 6, pp. 1–43, 2013.

[7] T. Suel and N. Memon, “Algorithms for delta compression and remote file synchronization,” 2002.

[8] I. Drago, E. Bocchi, M. Mellia, H. Slatman, and A. Pras, “Benchmarking personal cloud storage,” in Proceedings of the 2013 conference on Internet measurement conference. ACM, 2013, pp. 205–212.

[9] I. Drago, M. Mellia, M. M Munafo, A. Sperotto, R. Sadre, and A. Pras, “Inside dropbox: understanding personal cloud storage services,” in Proceedings of the 2012 ACM conference on Internet measurement conference. ACM, 2012, pp. 481–494.

[10] U. Manber et al., “Finding similar files in a large file system.” in Usenix Winter, vol. 94, 1994, pp. 1–10.

[11] P. Mahajan, S. Setty, S. Lee, A. Clement, L. Alvisi, M. Dahlin, and M. Walfish, “Depot: Cloud storage with minimal trust,” ACM Transactions on Computer Systems (TOCS), vol. 29, no. 4, p. 12, 2011.

[12] A. J. Feldman, W. P. Zeller, M. J. Freedman, and E. W. Felten, “Sporc: Group collaboration using untrusted cloud resources.” in OSDI, vol. 10, 2010, pp. 337–350.

[13] F. Dabek, M. F. Kaashoek, D. Karger, R. Morris, and I. Stoica, “Wide-area cooperative storage with cfs,” in ACM SIGOPS Operating Systems Review, vol. 35, no. 5. ACM, 2001, pp. 202–215.

[14] L. P. Cox and B. D. Noble, “Samsara: Honor among thieves in peer-to-peer storage,” ACM SIGOPS Operating Systems Review, vol. 37, no. 5, pp. 120–132, 2003. [15] H. Zhuang, R. Rahman, and K. Aberer, “Decentralizing the cloud: How can small data centers cooperate?” in Peer-to-Peer Computing (P2P), 14-th IEEE International Conference on. Ieee, 2014, pp. 1–10.