INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 3 ISSUE 6 (2017) PAGES 2541 - 2545
RECEIVED : 01.11.2017 PUBLISHED : 04.12.2017 Dec 04, 2017

Secure Data De-Duplication With Dynamic Ownership Management In Cloud Storage

[1]A. Sumathi, PG Scholar, Department Of Computer Science And Engineering, Maha Barathi Engineering College, Chinna Salem, Villupuram,

[2]S.Pavithra Devi, Assistant Professor, Department Of Computer Science And Engineering, Maha Barathi Engineering College, Chinna Salem, Villupuram.

Abstract:

Data de-duplication is used in cloud storage to save bandwidth and reduce the storage space by keeping only one copy of same data. But it raises problems involving data ownership and security when multiple users upload the same data to cloud storage. Since encryption preserves privacy, yet its randomization property hampers de-duplication. Hence, there is a need of secure data deduplication scheme to prevent unauthorized access and data leakage. In recent times, a number of de-duplication schemes have been proposed to solve this problem. However, many systems suffer from security flaws because they do not reflect the dynamic changes in the ownership of outsourced data. In this paper, we review several deduplication techniques over encrypted data to achieve secure and efficient cloud storage service. Furthermore, proposed scheme uses RCE and group key management mechanism to ensure that only authorized access to the shared data is possible, which is considered to be the most important challenge for secure and efficient cloud storage service in the environment where ownership changes dynamically.

Keywords: De-duplication, cloud storage, encryption, proof-of-ownership.

1. INTRODUCTION

Cloud Computing is a widespread term used in today's world. It delivers infinite space for storage, readiness, user-friendliness from anywhere, anytime to entities. Now-a-day's number of users and their data in the cloud is continuously growing with higher memory space and upload bandwidth. Data de-duplication used in cloud storage providers to resolve these overheads. De- duplication is a process of removing multiple copies of same data, to reduce the storage space and save bandwidth. But when same data outsourced by users to cloud storage some challenges are arises on data ownership and security for sensitive data. Today's cloud storage services like Dropbox and Google Drive etc. use a de-duplication scheme to save the network bandwidth and the storage cost. As data owners worried about their private data, they may encrypt their data before uploading in order to keep data privacy from illegal outside adversaries, as well as from the cloud service provider.

As concern with authorized access and security, there are many encryption schemes proposed. De-duplication scheme takes benefit of data similarity to find the same data and scale down the storage space. In contrast, encryption algorithms randomized the encrypted files to make cipher-text same from theoretically random data. Encryption of the same data by dissimilar users with different encryption keys results in different ciphertexts, which makes it hard for the cloud server to decide whether the plain data are the same and de-duplicate them. Hence, traditional encryption makes de-duplication impossible for above reasons.

2. RELATED WORK

The simplest implementation of traditional encryption can define as follows: Consider users A and B, encrypts the same file M under their secret keys SKA and SKB and stores corresponding cipher-text CA and CB. Then, further problems arise: First, how can the cloud server sense that the underlying file M is similar, and second is even if it can notice this, how can it allow both users to recover the stored data, based on their distinct secret keys? One simple way out is to let on each client to encrypt the file with the public key of the cloud storage server. Then, the server is capable to de-duplicate the identified data by decrypting it with its private key pair. Still, this solution grants access to the cloud storage server to get the outsourced plain data, which may break up the privacy of the data if the cloud server cannot be fully trusted. Convergent encryption plays the vital role in data de- duplication and overcomes the drawback which discussed above. A convergent encryption algorithm works as follows: Firstly, it takes an input file and encrypts them with its hash value as an encryption key. Then, the ciphertext is given to the cloud server and user keeps the encryption key. As convergent encryption is deterministic, every time similar files encrypted into similar cipher-text irrespective of who encrypts them.Hence, the cloud server can do de-duplication over the generated ciphertext. Then all data owners can download the ciphertext and decrypt it later as they have the same encryption key for the file. But convergent encryption has security weakness concern with tag consistency and ownership revocation.

3. LITERATURE SURVEY

In order to keep data privacy against inside cloud server as well as outside challengers, users may want their data encrypted. However, conventional encryption under different users' keys makes cross-user de-duplication impossible, since the cloud server would always see different ciphertexts, even if the data are the same, regardless of whether the encryption algorithm is deterministic.Douceur [2] introduces Convergent Encryption, which is the promising solution to this problem.

Bellare [3] introduces an idea of message-locked encryption (MLE), with its security approach to solving the problem of CE. He also proposed randomized convergent encryption (RCE) as one application of MLE which provides a technique to achieve secure de- duplication. In RCE, initial uploader encrypts a message using a random encryption key and it results into a ciphertext refer as C1. This message encryption key is again encrypted along with a key encrypted key (KEK) which is derived from the message by using hash function and results into a ciphertext refer as C2. Here message tag is generated from the KEK.

Xu[4] proposes a leakage-resilient de-duplication scheme to solve the data integrity issue. It addressed a vital security concern in cross-user client-side de-duplication of encrypted files in the cloud storage: privacy of users' sensitive files against both outside challengers and the honestbut-curious cloud storage server in the bounded leakage model.

Instead of encrypting the convergent keys on a per-user basis, Dekey builds secret shares on the original convergent keys (that are in plain) and assigns the shares over various KM- CSPs. If many users share the identical block, they can access the same corresponding convergent key. This significantly decrease the storage overhead for convergent keys. In addition, this method provides fault tolerance and allows the convergent keys to remains accessible even if any subset of KM-CSPs fails.

4. PROPOSED SYSTEM

In the existing deduplication system, each user is issued a set of privileges during system initialization. Each file uploaded to the cloud is also bounded by a set of privileges to specify which kind of users is allowed to perform the duplicate check and access the files. Before submitting his duplicate check request for a file, the user needs to take this file and his own privileges as inputs. The user is able to find a duplicate for this if and only if there is a copy of this file and a matched privilege stored in cloud. In the proposing system, we .eliminating duplicate copies of repeating data and has been widely used in cloud storage to reduce the amount of storage space and save bandwidth.

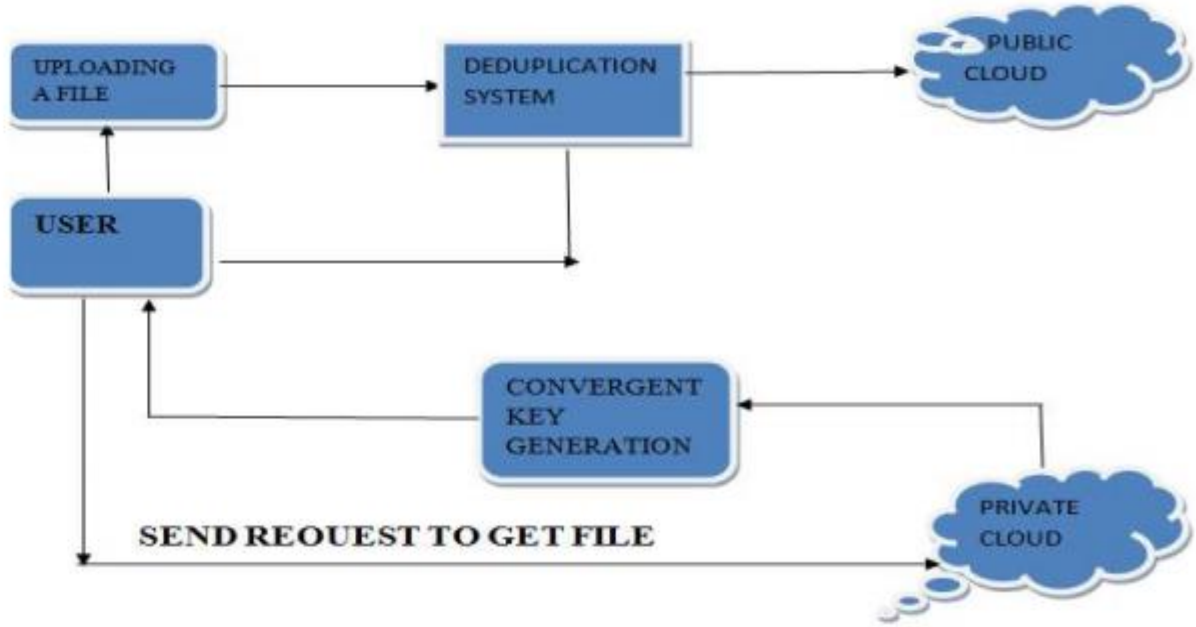
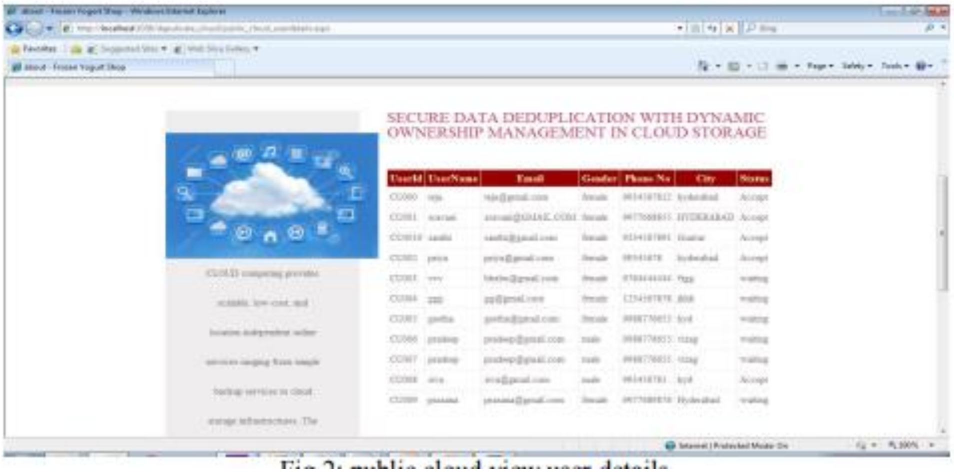


Fig.1.System Architecture

To protect the privacy of sensitive data while supporting deduplication, the convergent encryption technique has been proposed to encrypt the data before outsourcing .To better protect data security, this paper makes the first attempt to formally address the problem of authorized data deduplication Whenever someone wants to give information or take information from cloud they have to take permission i.e. authentication is done. If not a member they have to register first. Then the user will request private cloud to get a file token .Private cloud will issue file token and Convergent key generation takes place. With that key user will upload a file to the public cloud, then there will be a deduplication system to check whether the file already exist or not. If the file already exists then the file will not upload in public cloud.

5. ANALYSIS

A user derives a convergent key from each original data copy and encrypts the data copy with the convergent key. The key generation algorithm that maps a data copy to a convergent key.


About - Frozen Yogurt Shop - Windows Internet Explorer
Favorites
Suggested Sites
Web Slice Gallery
About - Frozen Yogurt Shop
Page
Safety
Tools
Internet | Protected Mode: On
100%
Fig 2: public cloud view user details

SECURE DATA DEDUPLICATION WITH DYNAMIC OWNERSHIP MANAGEMENT IN CLOUD STORAGE

CLOUD computing provides scalable, low-cost, and location-independent online services ranging from simple backup services to cloud storage infrastructures. The

UserId	UserName	Email	Gender	Phone No	City	Status
CC000	raja	raja@gmail.com	female	[illegible]	hyderabad	Accept
CC001	sravani	sravani@GMAIL.COM	female	[illegible]	HYDERABAD	Accept
CC0010	[illegible]	[illegible]@gmail.com	female	[illegible]	[illegible]	Accept
CC002	priya	priya@gmail.com	female	[illegible]	hyderabad	Accept
CC003	vvv	[illegible]@gmail.com	female	[illegible]	[illegible]	waiting
CC004	ggg	gg@gmail.com	female	[illegible]	ddd	waiting
CC005	geetha	geetha@gmail.com	female	[illegible]	hyd	waiting
CC006	pradeep	pradeep@gmail.com	male	[illegible]	vizag	waiting
CC007	pradeep	pradeep@gmail.com	male	[illegible]	vizag	waiting
CC008	siva	siva@gmail.com	male	[illegible]	hyd	Accept
CC009	prasanna	prasanna@gmail.com	female	[illegible]	Hyderabad	waiting

Fig.2.Cloud Viewer

The symmetric encryption algorithm that takes both the convergent key and the data copy as inputs and then outputs a cipher text. The decryption algorithm that takes both the cipher text and the convergent key as inputs and then outputs the original data copy and the tag generation algorithm that maps the original data copy and outputs a tag.

CONCLUSION

In this paper, we have reviewed different data deduplication techniques over encrypted data that are used in the cloud computing for secure data storage. Traditional encryption makes deduplication impossible because of the randomization property of encryption. Recently, several deduplication schemes are proposed to solve this issue by allowing each owner to share the same encryption key for the same data. Convergent encryption has different encryption variants for secure deduplication which was formalized as MLE later in. Though, CE suffers from security flaws with regard to tag consistency and ownership revocation. The schemes MLE and LR areproposed to recover the drawback of CE but still these schemes are insecure in the setting of PoW and Dynamic Ownership Management among the data owners. Furthermore, many schemes could not achieve secure access control under dynamic environment. Hence, not much work has yet been done to address dynamic ownership management and its related security problem. Thus the proposed scheme ensures that only authorized access to the shared data is possible, which is considered to be the most important challenge for efficient and secure cloud storage services in the environment where ownership changes dynamically.

REFERENCES

[1] J. Li, X. Chen, X. Huang, S. Tang, Y. Xiang, M. Hassan, and A. Alelaiwi, “Secure Distributed Deduplication Systems withImproved Reliability,” IEEE Transactions on Computer, Vol. 64, No. 2, pp. 3569–3579, 2015.

[2] R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M.Theimer, “Reclaiming space from duplicate files in a server less distributed file system,” Proc. International Conference on Distributed Computing Systems (ICDCS), pp. 617–624, 2002.

[3] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-locked encryption and secure deduplication,” Proc. Eurocrypt 2013, LNCS 7881, pp. 296–312, 2013. Cryptology ePrint Archive, Report 2012/631, 2012.

[4] Xu, E. Chang, and J. Zhou, “Leakage-resilient client-side deduplication of encrypted data in cloud storage,” ePrint, IACR, http://eprint.iacr.org/2011/538.

[5] Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou, “Secure deduplication with efficient and reliable convergent key management,” IEEE Transactions on Parallel and Distributed Systems, Vol. 25, No. 6, 2014.

[6] J. Li, Y. K. Li, X. Chen, P. Lee, and W. Lou, “A hybrid cloud approach for secure authorized deduplication,” IEEE Transactions on Parallel and Distributed Systems, Vol. 26, No. 5, pp. 1206–1216, 2015.

[7] X. Jin, L. Wei, M.Yu, N.Yu and J. Sun, “Anonymous deduplication of encrypted data with proof of ownership in cloud storage,” Proc. IEEE Conf. Communications in China (ICCC), pp.224-229, 2013.