International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

A Cloud Storage System for Preserving Privacy and Data Integrity of User

M. S. Tore[1], S. K. Sonkar[2]

[1, 2] Savitribai Phule Pune University, AVCOE Sangamner

Abstract: ***Cloud computing offers well-known services for storing user data and it gives attention towards a broad set of policies, technologies and controls deployed to facilitate security for applications and data. As the more and more businesses using the cloud, security is becoming main problem in the cloud environment. It is very important that companies work with partners who knows best practices of cloud security and which provides transparency for their solutions. Many security solutions today rely on the authentication for providing security but it didn't solve the privacy issues while sharing data over the cloud storage. Users data access request itself may reveal users private information no matter his request approved or not. So this become the important in data sharing in the cloud computing. In this paper we proposed a system which addresses the above mentioned issue. The proposed system uses the concept of data anonymity for anonymous matching of data access request. Our system also provides the data auditing functionality to protect the integrity of users shared data.***

Keywords: Cloud computing, data anonymity, user privacy, data integrity, security.

1. Introduction

Integration of web services and data centers offered by cloud computing in distributed computing style. Major companies like Amazon, Google, Microsoft, Yahoo, Salseforce and others offers cloud services to users. At the initial stage Amazon provided an architecture of cloud based services and after this many new models and enhancement has been proposed for cloud architecture.

Today there are many techniques available for storing data over a cloud server so that client can be guaranteed in terms of CIA triad, i.e confidentiality, integrity and availability of data over cloud servers. Confidentiality refers to the fact that information is secured from the unauthorized entities. It can be ensure by various cryptographic techniques. Availability means the user can use his own information stored over cloud server everywhere and anytime. Integrity refers to the fact that actual data cannot be altered by any external entities.

Cloud technology can increase availability with the help of internet enabled access, but in this case client is restricted to timely and robust provision of resources. Availability is influenced by the architecture of a cloud and capacity of cloud to store data of cloud provider.

Figure. 1 [1] shows the cloud environment which includes three elements. First is the cloud user who has the large amount of data files to be stored over cloud, second is the cloud server which is managed by the cloud service provider which is able to provide storage services and has significant storage space and computation resources and the last one is the Trusted Third Party (TPP) who is expert and has the capabilities that cloud user does not has. It is trusted to evaluate cloud storage service reliability on behalf of user upon user's request.

Users store their data remotely and take benefits of pull based top quality applications and services from a shared pool of configurable computing resources using cloud storage, without any burden of local data maintenance and storage. The other benefit of using a cloud services user can easily share their data with other users for getting more and more profits. So data sharing over the cloud computing is an important functionality in cloud computing. Though data sharing becomes more important in cloud computing, it also raises some security and privacy issues in cloud data sharing environment. User's data access request can itself reveal users privacy no matter whether he get the permission for accessing data. Also user storing data over cloud no longer has physical possession of data, it makes the protection of data integrity a difficult task. So there is a need to design a system which can protect user privacy and data integrity and also can be able to detect the changes made to the shared data over a cloud.

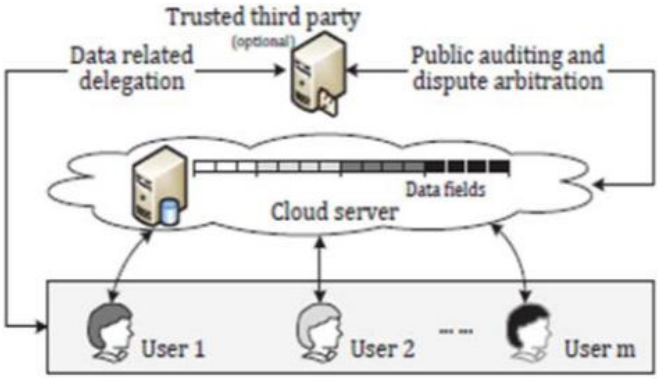

Figure 1: Cloud Storage Environment

So taking the need of cloud computing in shared tenancy environment we are designing a system which address the above mentioned security issues related to user's privacy and data integrity, our proposed system protect user privacy by using anonymous request matching mechanism when user requests other user's data over cloud storage. To share data with more than one user proxy re-encryption [] is provided by cloud server. Attribute based encryption is used so that user can only access its own data fields. Our system also checks user data for its integrity by acting as a third party

Licensed Under Creative Commons Attribution CC BY

auditor (TPA). So in short our system will provide a secure way for sharing data over cloud storage.

2. Related Work

Liu et al. [1] proposed a shared authority based privacy preserving protocol which allows cloud users to get shared access authority by anonymous access request matching mechanism. It also adopts attribute based access control so that user can only access its own data fields. Cloud server uses proxy re-encryption to provide data sharing among multiple users.

Cong Wang et al. [2] proposed a secure cloud storage system supports a privacy preserving public auditing. Third Party Auditor (TPA) is introduced to check user's data for its integrity. Proposed TPA can also able to perform audits for multiple users simultaneously and efficiently. Homomorphic linear authentication scheme (HLA) and Message authentication scheme (MAC) schemes are used to check data integrity.

Dunning et al. [3] proposed an anonymous ID assignment based data sharing algorithm (AIDA) for multiparty oriented cloud and distributed computing systems. In AIDA secure sum data mining operations are used to design an integer data sharing algorithm. It also adopts a variable and unbounded number of iterations for anonymous assignment. It uses Sturm's theorem and Newton's identities for data mining operations. Algorithms scalability is enhanced by using distributed solutions of certain polynomials over finite fields. To determine statistics of the required number of iterations Markov chain representations are used.

Liu et al. [4] proposed a MONA (Multi-owner Data Sharing Secure Scheme) for dynamic groups in the cloud applications. With the help of MONA, user is able to share its data securely with other users via untrusted cloud server and also support dynamic group interactions efficiently. In this scheme of data sharing, without pre-contacting with owner of data, a new granted user is able to decrypt data files. A user revocation is done with the help of revocation list. In the process of user revocation secret keys of the remaining users are not updated in revocation list. Any user in the group can utilize cloud resources anonymously. This is confirmed by applying access control. For dispute arbitration only group manager can reveal true identity of data owner. Here it indicates that the encryption computation cost and storage overhead is not related to the amount of users.

Wang et al. [5] proposed a distributed storage integrity auditing mechanism to enhance secure and dependable storage services in cloud computing. For this purpose the Homomorphic token and distributed erasure coded data is introduced. The scheme allows users to take audit of the cloud storage. For this the communication and computation cast is made low. The auditing results guarantee the correctness of cloud storage. This scheme also supports dynamic outsourced data operations. It shows that the scheme is able to recover quickly against malicious data modification attacks, server colluding attacks and Byzantine failure. To improve the weakness of symmetric key cryptosystem in public clouds.

Nabeel et al. [6] proposed a broadcast group key management (BGKM) and it realizes that user need not realize public key cryptography , and can dynamically derive the symmetric key during decryption. According to this, attribute based access control mechanism is designed so it can achieve that user can decrypt the data if and only if its attribute of identity satisfy the content provider's policies. For assigning secrets to users based on the identity attributes fine grained algorithm applies Access Control Vector (ACV) and allow users to generate symmetric keys based on their secrets and other public information. The BKGS has an advantage when there is need for adding/revoking users and updating access control policy.

In above works, various security related issues are addressed that can cause data sharing in cloud. However users access request related privacy and data integrity issues are not studied together. So here we are studying these issues together and designing a system which is able to preserve user's privacy while sharing data over cloud and also considering issue related to user's data integrity.

3. Problem Statement

"The aim is to develop a system which achieves shared access authority by matching anonymous access request by considering the user security and privacy and realize that user can only access its own data fields, share its data among multiple users by using proxy re-encryption. To preserve data integrity auditing functionality technique is applied"

4. Proposed System Model

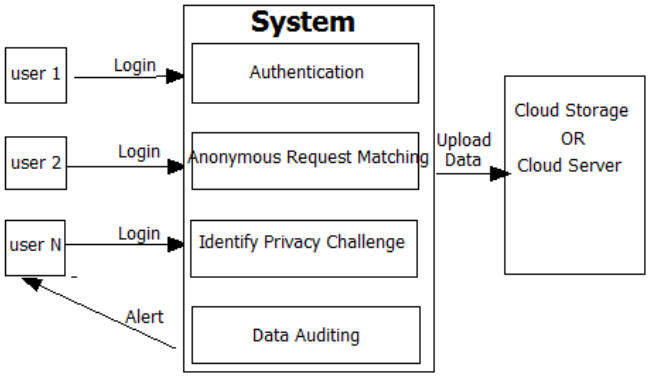

Figure 2: Proposed System Model

The architecture of proposed system for cloud storage as shown in figure 2. It consists of Users (Un), a cloud server (CS) and our proposed system. User can perform operations such as login to system, upload data to cloud server via System, request system for accessing other users' data etc. Cloud server (CS) stores user data in to a cloud storage it has. Main part of the architecture is system. It acts as a middleware between user and cloud server. It performs the functions for user such as authenticate the user, uploading user data over cloud server, anonymous access request

Licensed Under Creative Commons Attribution CC BY

matching for sharing the data, audit the user data for its integrity etc.

5. Design Goals

The proposed scheme consists of design goals that include anonymity, attribute based access control, proxy re-encryption and data auditing.

1. Anonymity: It guarantees that user get shared access authority of other users data without revealing the identity.
2. Attribute based access control: It ensures that each user can only get access to its own data fields.
3. Proxy re-encryption: It ensures that multiple user can share data among themselves.
4. Data auditing: it ensures that system checks user data for its correctness without retrieving the whole copy of the data and without the additional burden on the cloud server.

6. System Design and Methodology

The cloud storage system includes a cloud server CS and users {Ax} (x = {1,…,n}, n ϵ N^*). Consider A1 and A2 are two users having independent authorities on their own data fields. It indicates that user has access permission only for specific data fields stored by CS and it cannot exceed its authority access to obtain other users data fields. In this scenario we consider CS and {A1, A2} to describe our system.

Let BG = (q, g, h, G, G', e, H) be pairing group, in which q is a large prime, {G, G'} are of prime order q, G = <g> = <h> and H is a collision resistant hash function. The bilinear map e: G X G → G' satisfies the bilinear non-degenerate properties, i.e, for all g, h ϵ G and a, b ϵ Z_q^*, it turns out that e(g^a, h^b) = e$(g,h)^{ab}$ and e(g, h) ≠ 1. Meanwhile, e(g, h) can be efficiently obtain for all g,h ϵ G and it is a generator of G'. Let CS and Ax own the key pairs (pub_{cs}, pri_{cs}) and (pub_A, pri_A). Also cloud server CS is assigned with all users public keys $\{pub_{A_1}, pub_{A_2}, \ldots\ldots, pub_{A_x}\}$ and Ax is assigned with pub_{cs}. Public key is generated by $pub = g^{pri} \ (mod\ q)$ and the private key is generated by $pri\ \epsilon\ Z_q^*$ according to generator g.

Consider the following function,

$$F\left(REQ_{A_1}^{A_2}, \left(REQ_{A_2}^{A_1}\right)^T\right) = const \leftarrow Z_q \quad (1)$$

It describe the algebraic relation of $(REQ_{A_1}^{A_2}, REQ_{A_2}^{A_1})$ which are mutually inverse access requests challenged by (A_1, A_2) and *const* is a constant. Here F (.) is a collision resistant function for any randomized polynomial time algorithm A. There is a negligible function p (k) for a sufficiently large value k [1].

Prob [{(1, 1');(2, 2')} $\leftarrow A(1^k) : (1 \neq 1', 2 \neq 2') \wedge F\left(REQ_{A_1}^{A_2}, \left(REQ_{A'_2}^{A'_1}\right)^T\right) = const$] < p(k)

$REQ_{A_t}^{A_*}$ is a multidimensional Boolean vector in which only the $*^{th}$ element and the t^{th} element are 1 and other elements are 0. It shows the following[1] ,

1. $F\left(REQ_{A_1}^{A_2}, \left(REQ_{A_2}^{A_1}\right)^T\right) = F(2) = const$

   It shows that both A1 and A2 are interested in each other's data fields and two access requests are matched.

2. $F\left(REQ_{A_1}^{A_2}, \left(REQ_{A_2}^{A_1'}\right)^T\right) = F\left(REQ_{A_1}^{A_2'}, \left(REQ_{A_2}^{A_1}\right)^T\right) = F(1)$

   It shows that only one user (A1 or A2) is interested in the other's data fields and the access requests are not matched.

3. $F\left(REQ_{A_1}^{A_2'}, \left(REQ_{A_2}^{A_1'}\right)^T\right) = F(0)$

   It shows that no user (A1 or A2) is interested in each other's data fields and two access requests are not matched.

For checking data integrity, system uses the Message Authentication Code (MAC). The main idea is that system must calculate the MAC before uploading a data to a cloud server. At the time of data auditing MAC is calculated again and these two MACs are compare with each other. If the two MACs are same then there is no change in the data and the correctness of the data is proved. If the MACs are not equal then we can say that user data is tampered and it is informed to data owner.

Let F is a user file to be stored on to a cloud server and it is divided into 'n' number of blocks as F = $\{b_1, b_2, b_3, \ldots, b_n\}$ where n ϵ N*. MAC (.) is a function which computes the message authentication code.

Following are the steps that are performed for the Data Auditing:

1. Calculate the MAC before uploading each data block to a server, this can be done by using the CalcMAC function.
   $$MAC_{b_i old} = CalcMAC(b_i, k_s) \quad (2)$$
   where b_i is the data block of file F to be uploaded and k_s is a secret key used for generating MAC. After generating the MAC the data block is uploaded to a cloud server along with its calculated MAC.
2. In this step, actual data auditing is performed on the cloud server. First it is checked that cloud server has retain the whole copy of data block b_i . If this is not the case then data auditing is not performed by the system. MAC is calculated in following way,
   $$MAC_{b_i new} = CalcMAC(b_i, k_s) \quad (3)$$
3. This step checks that two MACs are same or not. For doing this the Boolean function named *compare(.)* is used as follows,
   Compare ($MAC_{b_i old}$, $MAC_{b_i new}$) (4)

If this function returns true then it indicate that both the MACs are same and data is not altered and if function return false then it indicate that MACs are different and data is altered.

Licensed Under Creative Commons Attribution CC BY

7. Experimental Setup

We are developing our system in C#.net and visual studio 2010. We used Blowfish algorithm for Encryption and Decryption. Also we used a K-Anonymity algorithm for sending anonymous data sharing request. We have Dropbox API for storing data over cloud and used SOAP for cloud linking. We also used PRNG (Pseudo Random Number Generation) algorithm for generating the keys required for Encryption.

8. Conclusion

In this paper, we have described a system which enables user to share their data securely without revealing their privacy. A system also audits the user data store over the cloud and preserve user's data integrity. Overall our system comprises user privacy and data integrity related issue while sharing data in a cloud environment and provides a secure way for for sharing data over the cloud. So our system guarantees that user's privacy and data integrity will be preserve. Our system will be applicable where there is use of cloud computing.

References

[1] H. Liu, H. Ning, Q. Xiong and L.T. Yang, "Shared Authority Based Privacy Preserving Authentication Protocol in Cloud Computing," IEEE Transactions on PARALLEL AND DISTRIBUTED SYSTEMS VOL:PP NO:99 YEAR 2014.

[2] C. Wang, S.S.M Chow, Q. wang , K Ren and W. Lou, *"Privacy-Preserving Public Auditing for Secure Cloud Storage,"* IEEE Transactions on COMPUTERS, VOL. 62, NO.2, FEBRUARY 2013.

[3] L. A. Dunning and R. Kresman, *Privacy Preserving Data Sharing With Anonymous ID Assignment,"* IEEE Transactions on Information forensics and Security, vol. 8, no. 2, pp. 402-413, 2013.

[4] X. Liu, Y. Zhang, B. Wang, and J. Yan, *"Mona: Secure Multi-Owner Data Sharing for Dynamic Groups in the Cloud,"* IEEE Transactions on Parallel and Distributed systems.[online]ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=6374615, 2012.

[5] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, *Toward Secure and Dependable Storage Services in Cloud Computing,"* IEEE Transactions on Services Computing, Vol.5, no. 2, pp. 220-232, 2012.

[6] M. Nabeel, N. Shang and E. Bertino, "Privacy reserving Policy Based Content Sharing in Public Clouds," *IEEE Transactions on Knowledge and Data Engineering*, [online] ieexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6298891, 2012.

Author Profile

Mr. Manoj S. Tore received the B.E. degree in In Computer Engineering from University of Pune, in 2012. Currently he is pursuing Master's degree in Computer Engineering from Amrutvahini college of Engineering, Sangamner under University of Pune. His areas of interest are network Security and cloud computing. He is currently working in the field of Network security and Cloud computing.

Prof. S. K. Sonkar received the Master degree in Computer science and Engineering from SRTMU Nanded. He is currently pursuing the Ph.D. degree in computer science from University of Pune. He is presently working as Assistant Professor in Dept. of Computer Engineering in Amrutvahini college of Engineering, Sangamner, India. His current research interests include network security and Cloud Computing.

Licensed Under Creative Commons Attribution CC BY