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

Privacy Preserved Public Auditing for Distributed Data

Anes P. A.[1], Neethu Francis[2]

[1]M.Tech Scholar, KMP College of Engineering, Cherukunnam, Perumbavur, India

[2]Assistant Professor, Department of Computer Science & Engineering, KMP College of Engineering & Technology, Perumbavur, India

Abstract: ***With distributed services strategy, it is commonplace for data to be not only stored in the server, but also shared across multiple users. However, public auditing for such shared data while preserving identity privacy remains to be an open challenge. In this paper, propose a privacy-preserved auditing on shared data stored in the distributed storage. In particular, exploit cryptographic signatures to compute the verification information needed to audit the integrity of shared data. With this mechanism, the identity of the signer in shared data is kept private from a third party auditor (TPA), who is still able to publicly verify the integrity of shared data. Experimental results can demonstrate the effectiveness and efficiency of the proposed mechanism when auditing shared data.***

Keywords: Public auditing, Shared data, Privacy-preserving, Distributed system, Trusted TPA, Data anonymization

1. Introduction

Distributed Computing [1] deals with distributed systems, which is a software system with networked computers communicate and coordinate their actions by passing messages. The software system components interact with each other in order to achieve a common goal. Grid computing [2] is a special type of distributed computing with non-interactive workloads that involve a large number of files. Distributed computing is just computing orchestrated between two or more computers. Cloud computing [3] is distributed computing, but a specialized form. The cloud computing service is either directly or indirectly, if use Amazon or Google we are directly storing in the cloud, using Twitter is an example of using cloud indirectly, as Twitter stores all tweets into the cloud.

Centralized computing systems [4], for example IBM Mainframes, controls all the peripherals and performs complex computations. However, centralized computing systems were ineffective and costly deal in huge volumes of transactional data and rendering support for tons of online users concurrently. Examples of distributed computing systems include: World Wide Web, Facebook, Hadoop's Distributed File System (HDFS), ATM, Google bots, Google web server, Indexing server etc...

The integrity of data in cloud or distributed storage is subject to scrutiny. To protect the integrity of data in that situation, it is best to perform public auditing by introducing a Third party auditor (TPA) [5], is designed to check the correctness of data stored in an untrusted server. The TPA is one who having the capabilities of doing the cloud or distributed operations on user's request. Third party auditor (TPA) also with distributed nature. TPA stores the users file related tokens and encrypted data and compares the received challenge results from different users (or servers). After verifying the file the TPA send the status of the file to the requested user. Here providing a trusted auditing mechanism why means the TPA only see the user file's encrypted text and token keys, it's enough to check or verify the storage correctness of the file.

TPA is one of the servers which are connected to the distributed system, the purpose of selecting TPA is that if the user doesn't have time to verify the privacy of data in cloud then they can select the TPA and it will provide the result based on a number of trials. It used a 256 bit AES algorithm for encryption/decryption purpose.

The first provable data possession (PDP) mechanism [6] to perform public auditing is designed to check the correctness of data. The proposed system provides public auditing along with data privacy and identity privacy. The distributed system mainly categorized under three heads viz. distributed information systems, distributed pervasive systems and distributed computing systems. The system in proposed paper refers distributed information system which uses communication models like RMI and RPC. Distributed pervasive systems include embedded computer devices such as portable ECG Monitors, wireless cameras etc... In distributed computing systems the computers connected within a network communicate through message passing to keep track of their actions. The proposed system achieves batch auditing where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA in a privacy-preserving manner. Our work is concentrated in the creation of a simulation in the field of distributed data storage privacy in cloud computing which do not normally provided by a typical cloud service provider (CSP) or service provider. The storage privacy checking is a trial and error strategy to identify the privacy measure of the cloud or distributed system with the help of a special user (TPA) without a resource centric direct communication between cloud server and normal user.

2. Problem Statement

People put their data on cloud service providers (CSP) or distributed server systems in belief that other users will not retrieve their privacy of data. In reality, the privacy of data may leak through hacker's intrusion or some unauthorized

actions. These problems could occur challenges to secure people's data privacy. Personal important data stored in the Cloud may contain banking information, passwords, important notes, and other important data that could be used. In this situation, an intrusive mechanism which can try to access the legitimate information on the request of legitimate user is an enhancing feature. The TPA can use any strategy of mechanism which reveals the inherent weakness of the cloud server on behalf of user. Mean time, the semi-trusted TPA must not leak any information regarding the data or identity. The working strategy of TPA is ethical in nature for futuristic security considerations.

2.1 System Model

As illustrated in Figure 1, the work in this paper involves three parties: the distributed (cloud) server, the third party auditor (TPA) and users. There are two types of users in a group: the original user and a number of group users. The original user and group users are both members of the group. Group members are allowed to view and download the data created by the original user based on access control policies. The third party auditor is able to verify the integrity of shared data in the cloud server on behalf of group members.

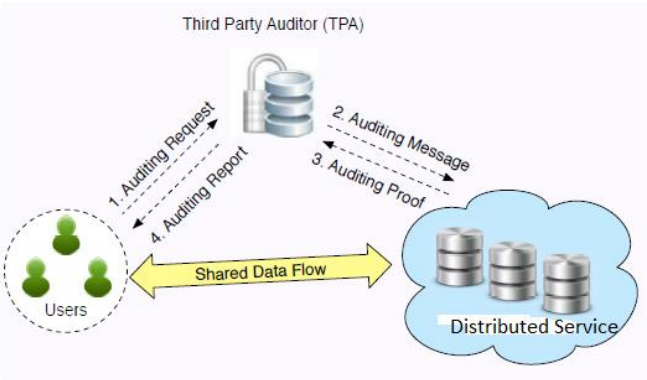

Figure 1: System model includes the distributed server, the third party auditor and users.

The enhancement in this paper allows adding dynamic group members. When the user (either the original user or a group user) wishes to check the integrity can request it to the third party auditor. After receiving the auditing request, the TPA generates an auditing message and generates a result based on the trial and error. The TPA gets the result and he can communicate it to the requested user with any means (not included in the paper). There is a passkey to protect the uploaded data file (in .txt format) and a token (which randomly generated based on read/write permission) which maps the permission on file.

2.2 Design Objectives

To enable the TPA efficiently and securely verify shared data for a group of users, the system should be designed to achieve properties: (1) Public Auditing: The third party auditor is able to publicly verify the privacy of shared data for a group of users without knowing the exact data. (2) Unforgeability: Only a user in the group can generate valid verification information on shared data. (3) Identity Privacy: During auditing, the TPA does not leak the identity of the requested party.

2.3 Data Anonymization

Anonymization is a technique that can use to increase the security of data in a distributed system while still allowing the data to be analyzed and used. The anonymization of data in file (through encryption) is done through AES (Advanced Encryption Standard) [7] algorithm, which is a symmetric one. Due to its symmetric nature it's faster than asymmetric algorithms such as RSA even though its complex to implement. Its security is based on cracking the AES key, which is infeasible in normal case.

Data anonymization is the key to preserve data privacy. The data can still be processed to gain useful information. Anonymized data can be stored in distributed environment without concern that other may capture the data. Later the results can be collected and mapped to the original data in a secure area.

3. Encryption Schemes

The AES algorithm is a block cipher with block length of 128 bits. It allows three key lengths: 128, 192 or 256 bits. Here use 256 bit as key length. Encryption consists of 14 rounds of processing. Except for the last round all rounds are identical. Unlike DES [8], the decryption algorithm differs substantially from the encryption algorithm, the same steps are used in algorithm but the order in which the steps are carried out is different.

The encryption strategy provides the data privacy from the semi-trusted TPA as well as any possible hacker or even from the CSP. The symmetric nature of algorithms like AES make the operation simple and light weight, avoids unnecessary complexity. The proposed system replaces the ring signature concept and uses efficient 256 bit AES crypto system. The enhancement due to this modification is the adding of dynamic users to the system without any modification in existing system as well as the system support batch auditing requests from multiple users.

4. Architecture Design

In this section mainly describe about the structure of proposed system in form of use case diagram. The system mainly consists of a distributed server (background service), user and a central TPA. In this system, multiple and independent users to connect directly to the service for both storage of data and can request auditing to TPA. The use case diagram for general functionality can be depicted as in Figure 2.

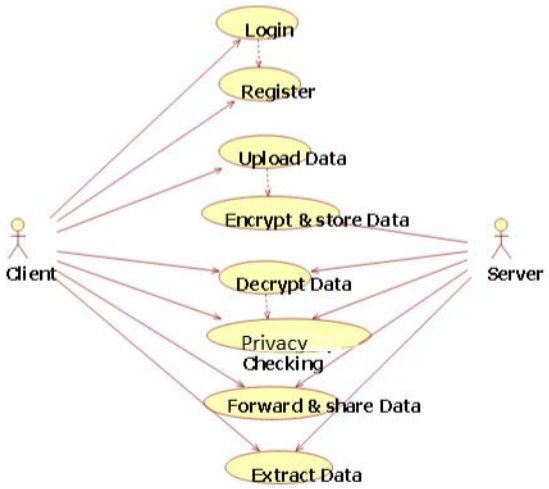

Figure 2: Architecture use case diagram

There is three distinct parts in this distributed system: user part (client), Server (Service or CSP) part and TPA. The events like Register, Login, Upload data, Forward & Share data, and Extract data are come under the part of user. The server part does the events like Encrypt & store data, Decrypt data. The TPA do the single event of Privacy checking, but there need an explicit from the part of user.

In the proposed system additionally there is an option to add dynamic users in the form of user registration. In the distributed storage the data file is stored as encrypted format to avoid the unauthorized modification.

5. Conclusion

The proposed system provides an innovative architecture that guarantees confidentiality of data stored in public distributed system. The main advantages are:

- Public auditability of shared data
- Privacy-preserving
- Data confidentiality
- Light weight
- Concurrent audit service

In general a public auditing consists of four procedures

- KeyGen - key generation algorithm that is run by the user to setup the scheme
- SigGen - used by the user to generate verification metadata, which may consist of MAC (message authentication code), signatures or other information used for auditing
- GenProof - run by the distributed server to generate a proof of data storage correctness
- VerifyProof - run by the TPA to audit the proof from the distributed server

There is only two access control permissions used. The read permission is applicable to all allowed users and write permission is allowed only to selected members. The read permission allows seeing file content but write permission allow downloading the same file.

The proposed scheme which rely on encrypting the data using some encryption algorithm (AES) and make third party auditor store a message digest or encrypted copy of the same data that is stored with the service provider. The third party is used to resolve any kind of conflicts (auditing here) between service provider and client. The distributed services can be provided in both web browser based and/or a light weight desktop app. The proposed system concentrate in an application level system not web based. The TPA here concentrates only in privacy auditing trials, not check any integrity or provide data security. But the data security and confidentiality is achieved through encryption based data anonymization.

References

[1] Cong Wang and Kui Ren, Illinois Institute of Technology WenjingLou, Worcester Polytechnic Institute Jin Li, Illinois Institute of Technology "Toward Publicly Auditable Secure Cloud Data Storage Services". 0890-8044/10/2010 IEEE.

[2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards Secure and Dependable Storage Services in Cloud Computing," IEEE Trans. Service Computing, vol. 5, no. 2, 220-232, Apr.-June 2012.

[3] K. Yang and X. Jia, "Data Storage Auditing Service in Cloud Computing: Challenges, Methods and Opportunities," World Wide Web, vol. 15, no. 4, pp. 409-428, 2012.

[4] Maheswaran, Muthucumaru; Shoukat Ali; Howard Jay Siegel; Hensgen, Debra; Freund, Richard F "Dynamic Mapping of a Class of Independent Tasks onto Heterogeneous Computing Systems" Calhoun Institutional Archive of the Naval Postgraduate School , June 1999.

[5] Cong Wang , Sherman S.M. Chow ,Qian Wang , Kui Ren,Wenjing Lou "Privacy-Preserving Public Auditing for Secure Cloud Storage" Issue No.02 - Feb. (2013 vol.62) pp: 362-375.

[6] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," in Proc. ACM Conference on Computer and Communications Security (CCS), 2007, pp. 598–610.

[7] Fei Shao;Nanjing, China ; Zinan Chang ; Yi Zhang "AES Encryption Algorithm Based on the High Performance Computing of GPU" 28 Feb. 2010.

[8] Seung-Jo Han ; Dept. of Electron. Eng., Chosun Univ., South Korea; Heang-Soo Oh; Jongan Park "The improved data encryption standard (DES) algorithm" 25 Sep 1996.

Author Profile

Anes P A received the B.Tech (CS) degree from Ilahia College of Engineering and Technology, Muvattupuzha, India in 2006. During 2006-2013, worked as Software Programmer in Web Technologies. In 2013-2015 pursuing M.Tech Computer Science and Engineering Specialization in Cyber Security. His research interests include Ethical hacking, Network security and GUI based software development.

Neethu Francis received the B.Tech (IT) from Viswajyothi College of Engineering in 2012 and M.Tech (CS) from Christ University Faculty Of Engineering, Kengeri, Bangalore in 2014. From 2014 November, worked as Assistant Professor in KMP College of Engineering, Perumbavur, India. Her research interests include Wireless Sensor Network, Mobile Computing and Computational Intelligence.