ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 7, July 2015

Privacy Preserving Public Auditing & Recovery for Secure Cloud Storage

Priya Rupeja[1],

Prof. K.C. Waghmare[2]
M.E., Computer Engineering, P.I.C.T, Pune, India [1]
Professor, Department of Computer Engineering, P.I.C.T, Pune, India [2]

Abstract: Cloud computing becomes most popular in today's era due to its charming characteristics like on demand self-service, board access network, resource pooling, multi-tenancy etc. Though it offers many characteristics but also prone to certain security issues. So that it becomes an obstacle for cloud computing development. Users rely on cloud for its storages, services etc. As users store their data on cloud they lost their physical control on it. In order to know their storage status then user has to rely on cloud service provider (CSP) report. There are certain chances that CSP may behave unfaithfully. To avoid this third party auditor (TPA) is introduced which will audit users data on behalf of them and generate report which contain information about its data integrity. In this paper we elaborated the functionality of TPA from providing data integrity information to notification details of lost data and its recovery.

Keywords: Cloud Computing, Data storage, TPA, data integrity, recovery, privacy-preserving.

I. INTRODUCTION

Cloud computing has became trending technology. According to NIST it is defined as model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [10]. User rely on cloud for many reasons like storage, services etc. By storing their data into the cloud, the data owners can be relieved from the burden of data storage and maintenance so as to enjoy the on-demand high quality data storage service. However, the fact that data owners and cloud server are not in the same trusted domain may put the outsourced data at risk, as the cloud server may no longer be fully trusted [20]. As users no longer physically possess control on its own data. For their data integrity user has to rely on cloud service provider report. There are certain reasons from which CSP behave unfaithfully. For example suppose CSP may delete users unused data to save their cloud storage and also it may hide certain data incidents in order to maintain its reputation.

To check whether CSP giving correct report, for that user has to download all the uploaded data and perform manual checking. This will become too cumbersome for user as well as enterprise. To fully ensure the data integrity and save the cloud users' computation resources as well as online burden, it is of critical importance to enable public auditing service for cloud data storage, so that users may resort to an independent third party auditor (TPA) to audit the outsourced data when needed. The TPA, who has expertise and capabilities that users do not, can periodically check the integrity of all the data stored in the cloud on behalf of the users, which provides a much more easier and affordable way for the users to ensure their storage correctness in the cloud [1]. TPA will audit user's information whenever user wants and generate its data integrity and consistency report which will be useful for both user as well as CSP. There are certain probabilities of violation of user's privacy. Because it might happen that TPA may learn information about outsourced data. To avoid that many encrypted techniques where applied over outsourced data but also that was less useful because its size increases. Hence lead to more space and cost.

Encryption techniques used only to ensure that TPA will not learn any knowledge. It will maintain user's privacy. But there are certain chances that hackers may find key and will decrypt that data and perform malfunction on it. So that lost or modified data cannot be recovered back. Our proposed system will overcome all this limitations and provide recovery of that lost or modified file so as to maintain the data integrity and consistency of user. We have

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 7, July 2015

shown certain possible attacks on cloud and its recovery. The rest of paper is organized is as follows. Section II gives the information about related work of proposed system. Section III introduces motivation behind proposed system. Then provide information of existing system in section IV, followed by detailed scheme for proposed system in section V. Section VI is for experimental setup, followed by results in section VII. Finally section VIII will give concluding remark of our whole system.

II. RELATED WORK

This section will give brief idea about techniques applied on data to reduce security threads on outsourced information. This section gives detailed overview of all techniques used before.

Ateniese et al. are the first one who consider public auditability concept in their defined provable data possession (PDP) model for ensuring possession of data files on untrusted storages for cloud. In this they used the RSA based homomorphic linear authenticators for auditing stored data and suggests randomly sampling algorithm for a few blocks of the file. However, it supposed to do the public auditability in their technique demands the linear combination of sampled blocks exposed to external auditor. When it comes for actual use then it notifies that there are certain chances by which auditor will learn that information from that outsourced data and not fully safe for privacy preserving of users data [19]. In this when they applied encrypted techniques then its data size is increased which lead to high data storage and cost.

Juels et al. Provides idea about a proof of retrievability (PoR) model, where spot-checking and error-correcting codes are used to ensure both possession and retrievability of data files on remote archive service systems. In this paper they initially define a number of audit challenges that user can do. They did not consider public auditability. They used Merkle-tree construction where they construct binary tree for public PoRs and it is applicable only on encrypted data [8].

Cong et al. specifies four algorithmic steps which gives idea about data integrity and consistency but not mentioned an idea about lost file and its recovery [1]. Cong Wong et al. focus on cloud data storage security, which has always been an important aspect of quality of service. To ensure the correctness of user's data in the cloud, they proposed an effective and flexible distributed technique with two features, opposing to its predecessors. It will utilize the homomorphic token with distributed verification of erasure-coded data, which achieves the integration of storage correctness insurance and data error localization, i.e., the identification of misbehaving server(s). In that they only identify misbehaving servers but not provide the information about corrupted file [7].

P. Selvigrija, D. Sumithra provides the idea about public auditing which can be achieved using the Automatic Protocol Blocking for the secure cloud storage, which improves the efficiency of the user storage. Thus the 3-d password will improve the user level security in Cloud Server and the data level security will be effectively provided using the GCM based encryption and decryption algorithms. Thus the public auditing ensures that the data leakage and data loss will be reduced [3].

Shah et al. they proposed technique which will keep storage as online by providing encryption on that. Then it performed by number of pre-computed steps and symmetric keyed hashes on that encrypted data to auditor. After that auditor in turn verifies integrity of that data file and by server's possession of previously committed decryption key. The limitation of this technique is that it works only on encrypted data file and suffers from auditors statefulness which may lead online burden to users [16], [17].

Sathiskumar R, Dr. Jeberson Retnaraj in this paper they proposed Encryption and Proxy encryption algorithm to protect the privacy and integrity of outsourced data in cloud Environments. This paper provides an abstract view of different schemes proposed in recent past for cloud data security using third party auditor [5].

Ming Li et al. in this they took dataset of health record of patients and perform encryption on it. They used attribute base technique. In that they took each record and encrypt it. After that it stores on cloud which leads to high storage. They didn't mention anything about lost of files and nothing about its recovery [4].

Jin Li, Qian Wang et al. in this paper, they exploit edit distance to quantify keywords similarity and develop an advanced technique on constructing fuzzy keyword sets, which greatly reduces the storage and representation overheads. In that they only consider privacy preserving but not its recovery of outsourced data [21].

III. MOTIVATION

From all the above limitations there is need for the system which will overcome all this drawbacks. Hence this system is proposed which will notify corrupted block and provide its recovery. Even that report is useful for CSP too from that it will come to know about their security status.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 7, July 2015

Hence CSP will do certain control measure to avoid all that loses possible to their cloud storage. Proposed system simulated some possible attacks on cloud storage like man in middle, session hijacking, DOS etc.

IV. EXISTING SYSTEM

In existing system there are three main entities which will perform all functions. Those are user, CSP and TPA. Here in this Cloud server and CSP is considered as a single entity. In this system three main tasks are performed i.e.

1. Uploading information

In this step user will upload information on cloud.

2. Stored on cloud

In this step all information of user is stored on cloud server.

3. Auditing Request

Here user will request to TPA to audit their outsourced data stored on cloud and provide the report whether integrity is maintained or lost.

V. PROBLEM DEFINITION

To develop system which provides notification and recovery of lost or corrupted file or block of outsourced data stored on cloud.

VI. PROPOSED FRAMEWORK

In the proposed system it basically consists of three entities that are cloud service provider (CSP), third party auditor (TPA), and user.

Firstly user registers them on system by filling all their details; if user is valid then they will get email from CSP which contain password. Using that password user will login to their respective account. They can upload their data on it. Before uploading of data on cloud, that file is divided into number of parts or blocks on that hashing is calculated using MD5 algorithm then that hash values are called as metadata. After that this metadata along with original data is stored on cloud. If user wants to check its data integrity then he/she will request TPA. Then TPA will request CSP to recalculate hash on that stored metadata by generating proof on it. After that TPA will get original metadata and newly calculated metadata. TPA verifies that proof by comparing old with new metadata.

If both are same then integrity is maintained otherwise integrity is lost. TPA will compare byte by byte and generate report. If multiple users are there at same time then batch auditing is done on it. And comparison will done using XOR technique. If some block is lost or corrupted then that is also mentioned in that report. So if user wants its recovery then he/she will request to CSP for that. Initially while uploading of data we replicate user's data and store it on web browser. CSP has direct access to that replicated data. If user wants recovery of its lost data then again splitting is done on that replicated data then corresponding lost block or file is recovered back. We also simulate certain types of possible attacks on cloud storage.

System contains following steps:

STEP 1: User Registration

In this step user will register them to the system by providing or filling their valid information like name, email id, phone no and address etc.

STEP 2: Approved by CSP

According to user's respective information CSP will approve user and register them as a valid user by providing user a confirmation mail with password.

STEP 3: User Login

After receiving password from CSP user will register to their account and change its account setting.

STEP 4: Upload Information & storing replication to Web server

User will upload their text document onto the cloud server and meanwhile the same file will be stored in web server in directory structure format. That is as user logged in to the system their stored information is also reflected to web server.

STEP 5: Splitting of File

In this step, file will be splitted into number of parts.

STEP 6: Calculation of Hash

On that parts hash is calculated using MD5 hashing algorithm.

STEP 7: Storing original information and hash value to cloud server

In this step original text file along with that hash values as a metadata to that original file is stored on cloud server.

STEP 8: On Demand Auditing

Here in this TPA will audit user stored file only when user request them. Then TPA will send request to CSP to give metadata and also request him to recalculate hash value on that stored file.

STEP 9: Comparing of hash values

In this step TPA will check original hash and recalculated hash value byte by byte.

STEP 10: Generate report

In this step report will be generate by TPA if data integrity and consistency only when all bytes are matched

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 7, July 2015

properly if not then integrity is lost. This will also provide information to user about their lost or corrupted file.

STEP 11: Report forwarded to user and CSP

In this step user will come to know there current information status on cloud storage.

STEP 12: Recovery Of data

In this step if user want to recovery their lost file then user will request to CSP then CSP will access that directory structure perform splitting and over it and recovery that lost file.

The proposed algorithm basically contain six phases

a) Setup Phase
b) SigGen Phase
c) Audit Phase
d) GenProof
e) Recovery
f) Attacks

In our system implementation mainly four modules are there:

1. Login and Registration
2. Upload & download (Hash calculation)
3. Auditing (For single as well as for multiple users)
4. Data Recovery & Attacks

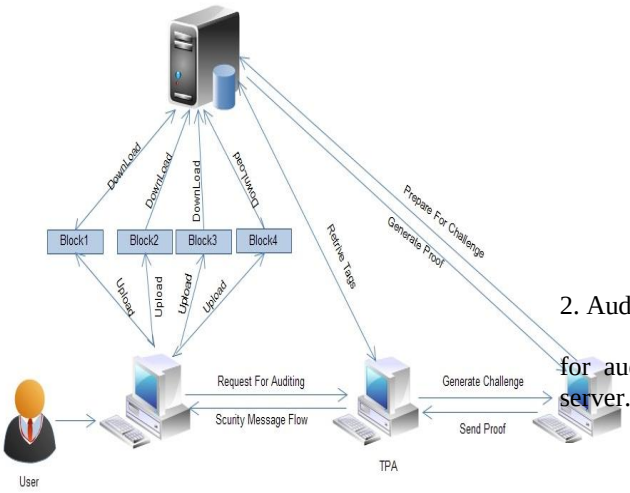

Fig. 1: System architecture of proposed system [2]

VII. EXPERIMENTAL SETUP

The experimental setup for the proposed system will require one virtual machine from Amazon cloud storage. On it we installed all software which used for our system implementation. User first register themselves on web browser if they are valid user then notification will send on them along with password of their account. Then user will store their information on cloud server. Even they can check their data integrity and also recovery their files by sending request to CSP. The particulars about platform and technology are as follows which are used to build proposed system:

Base Operating System: Window 7
Web Server: Tomcat server
Languages: java, javascript, xml
Database: MySQL 2008
Editor: Eclipse Luna
Deployed on Amazon cloud storage.

VIII. RESULT

Following below are the snap shots of GUI of our proposed system:

1. Upload text files:

In this user upload their data on cloud server without any storage limitation.

Snap 1: Upload file

2. Auditing Request:

In this user, whenever wants send its data to TPA for auditing and gets its actual report or status on cloud server.

Snap 2 : Auditing Request

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 2, Issue 7, July 2015

3. Recovery:

In this, user comes to know about its data status like which block of its respective file is modified and he/she will recover it by requesting to CSP.

Snap 3: Recovery of Files

IX. CONCLUSION

In existing system TPA will audit user's information on behalf of them and generate report which contains information about its data integrity. It will not provide any information about which data is lost. In this paper we tried to overcome limitations of present scenario by extending the functionalities of TPA. In this TPA will provide information about lost data. Even it also gives exact idea about which file is corrupted and from that also notifies which block is modified. It will provide recovery of that lost or corrupted file or block. It also gives idea about types of attacks possible on outsourced data of user and provides its recovery.

REFERENCES

[1]. Cong Wang, Sherman S.-M, Qian Wang, Kui Ren, Wenjing Lou,"Privacy-Preserving Public Auditing for Secure Cloud Storage", IEEE Trans. On Cloud Computing, March-2013

[2]. Priya Rupeja , Prof. Kalyani Waghmare, Privacy Preserving Public Auditing and Recovery using Backup & Restore Method for Secure Cloud Storage, in IJECS Volume 4 Issue 1 January, 2015 Page No.9929-9932.

[3]. P. Selvigrija, D. Sumithra, Public Auditing & Automatic Protocol Blocking with 3-D Password Authentication for Secure Cloud Storage, D. Sumithra et al, International Journal of Computer Science and Mobile Applications,Vol.2 Issue. 1, January- 2014, pg. 1-8

[4]. Ming Li, Shucheng Yu, Yao Zheng, Kui Ren, and Wenjing Lou, Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption, IEEE Trans on Parallel and Distributed Systems, Vol. 24, No. 1, Jan 2013.

[5]. Sathiskumar R, Dr. Jeberson Retnaraj ,Secure Privacy Preserving Public Auditing for Cloud storage, IJIRSET Volume 3, Special Issue 1, January 2014.

[6]. U. Jyothi K., Nagi Reddy, B. Ravi Prasad, Achieving Secure, Scalable, and Fine grained Data Access Control in Cloud Computing, International Journal Of Engineering And Computer Science ,Volume 2 Issue 8 August, 2013 Page No. 2440-2447

[7]. Cong Wang, Qian Wang, and Kui Ren, Wenjing Lou, Ensuring Data Storage Security in Cloud Computing

[8]. A. Juels and J. Burton S. Kaliski, Pors: Proofs of retrievability for large files, in Proc. of CCS07, Alexandria, VA, October 2007, pp. 584597.

[9]. Y. Dodis, S. P. Vadhan, and D. Wichs, Proofs of retrievability via hardness amplification, in TCC, 2009, pp. 109127.

[10]. P. Mell and T. Grance, Draft NIST working definition of cloud computing, Referenced on June. 3rd, 2009 Online at http://csrc.nist.gov/groups/SNS/cloud-computing/index. html, 2009.

[11]. Tejashree Paigude, Prof. T. A. Chavan, A survey on Privacy Preserving Public Auditing for Data Storage Security , in International Journal of Computer Trends and Technology- volume4 Issue3- 2013

[12]. G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, Scalable and efficient provable data possession, in Proc. Of Secure Comm 08, 2008, pp. 110.

[13]. Salve Bhagyashri1, Prof. Y.B.Gurav2, a Survey on Privacy-Preserving Techniques for Secure Cloud Storage, in IJCSMC, Vol. 3, Issue. 2, February 2014, pg.675 680

[14]. Anup Mathew, Survey Paper on Security & Privacy Issues in Cloud Storage Systems, in EECE 571B, TERM SURVEY PAPER, APRIL 2012

[15]. H. Shacham and B. Waters, Compact proofs of retrievability, in Proc. of Asia crypt 2008, vol. 5350, Dec 2008, pp. 90107.

[16]. M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, Auditing to keep online storage services honest, in Proc. Of HotOS07. Berkeley, CA, USA: USENIX Association, 2007, pp. 16

[17]. M. A. Shah, R. Swaminathan, and M. Baker, Privacy preserving audit and extraction of digital contents, Cryptology ePrint Archive, Report 2008/186, 2008

[18]. M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, Above the clouds: A berkeley view of cloud computing, University of California, Berkeley, Tech. Rep.

[19]. G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, Provable data possession at untrusted stores, in Proc. Of CCS07, Alexandria, VA, October 2007, pp. 598609.

[20]. D. Boneh, B. Lynn, and H. Shacham, Short signatures from the Weil pairing, J. Cryptology, vol. 17, no. 4, pp. 297319, 2004

Jin Li, Qian Wang, Cong Wang, Ning Cao, Kui Ren, and Wenjing Lou, Fuzzy Keyword Search over Encrypted Data in Cloud Computing , INFOCOM, 2010 Proceedings IEEE.