PRIVACY PRESERVING PUBLIC AUDITING SHARED DATA ON CLOUD BY USING TPA AND VISUAL CRYPTOGRAPHY

Jyoti Raykar, Snehal Shinde, Bhagyashree Chavan, Niyati Pandit

Abstract— To protect outsourced data in cloud storage against corruptions, adding fault tolerance to cloud storage together with data integrity checking and failure reparation becomes critical. Recently, regenerating codes have gained popularity due to their lower repair bandwidth while providing fault tolerance. Existing remote checking methods for regenerating-coded data only provide private auditing, requiring data owners to always stay online and handle auditing, as well as repairing, which is sometimes impractical.

We propose a third party auditing scheme for checking de-duplication and regenerating-code-based cloud storage. To solve the regeneration problem of failed authenticators in the absence of data owners, we introduce a web service, which is privileged to regenerate the authenticators, into the traditional third party auditing system model. We design a novel third party verifiable authenticator, which is generated by a couple of keys and can be regenerated using partial keys. Thus, our scheme can completely release data owners from online burden.

Index Terms—Data storage, Privacy-preserving, Public auditing, Cloud computing, Cryptographic protocol, Security, De-duplication, Shared data .

I. INTRODUCTION

To protect outsourced data in cloud storage against corruptions, adding fault tolerance to cloud storage together with data integrity checking and failure reparation becomes critical. Recently, regenerating codes have gained popularity due to their lower repair bandwidth while providing fault tolerance.We propose a third party auditing scheme for checking de- duplication and regenerating-code-based cloud storage. To solve the regeneration problem of failed authenticators in the absence of data owners, we introduce a web service, which is privileged to regenerate the authenticators, into the traditional third party auditing system model. Moreover, we design a novel third party verifiable authenticator, which is generated by a couple of keys and can be regenerated using partial keys. Thus, our scheme can completely release data owners from online burden.Cloud Computing Cloud computing, to put it simply, means "Internet Computing." The Internet is commonly visualized as clouds; hence the term "cloud computing" for computation done through the Internet. With Cloud Computing users can access database resources via the Internet from anywhere, for as long as they need, without worrying about any maintenance or management of actual resources. Besides, databases in cloud are very dynamic and scalable." Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources.

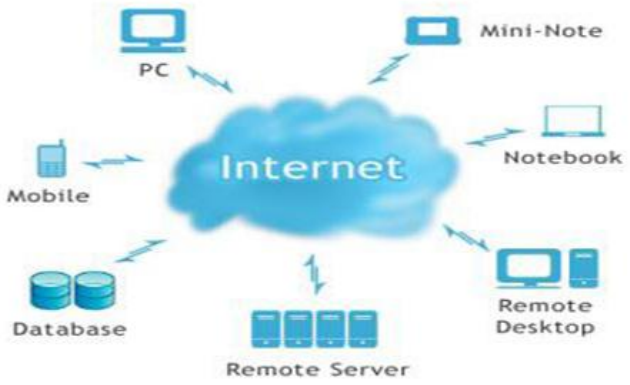
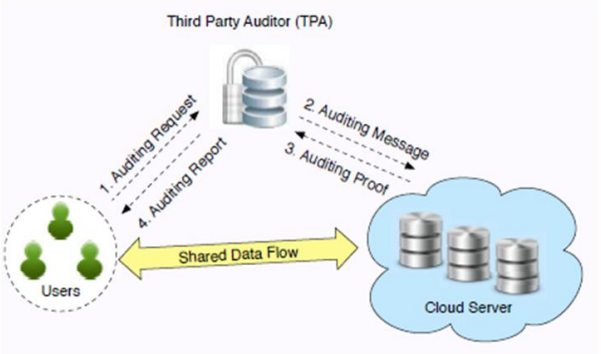


Fig 1: CLOUD COMPUTING

II. EXISTING SYSTEM

In existing system how to audit the integrity of the shared data in the cloud with **static** groups. It means the group is pre-defined before shared data is created in the cloud and the membership of users in the group is not changed during data sharing. The original user is responsible for deciding who is able to share her data before outsourcing data to the cloud. Another interesting problem is how to audit the integrity of shared data in the cloud with **dynamic** groups a new user can be added into the group and an existing group member can be revoked during data sharing while still preserving identity privacy.



Manuscript received Jan, 2016.
Niyati Pandit, *Computer Department, Savitribai Phule Pune University, Pune, India, 9561628982.*
Jyoti Raykar, *Computer Department, Savitribai Phule Pune University, Pune, India,99561272412.*
Snehal Shinde, *Computer Department,Savitribai Phule Pune University, Pune,India,7709916146*
Bhagyashree Chavan *Computer Department, Savitribai Phule Pune University, Pune, India,8600927852*

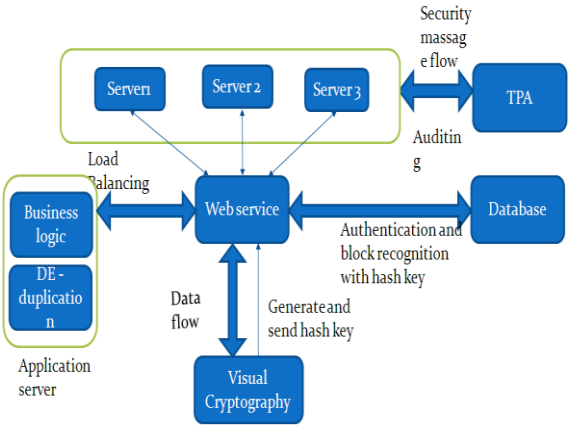
ISSN: 2278 – 1323

III. PROBLEM STATEMENT

To develop a web system that performs Data integrity checking and failure repairing on data being uploaded on cloud as a third party without disturbing the privacy of the user with de-duplication giving efficient bandwidth utilization.

IV. PROPOSED SYSTEM

In our system we are proposing to develop an application which increase the security provided by visual cryptography and Provides Third Party Auditing. Third party auditing scheme for checking regenerating-code-based cloud storage



V. PROCESS

In this system a user can upload files to the cloud using cloud services, the TPA provides security to both the cloud as well as the data being uploaded .At the web service the user data is split into number of parts and each part will be stored at different servers. The data will be stored in database initially, as the user upload a file. A hash key is generated at the user and also at database.

The database check whether there already exist a copy of user data,if it exist then only hash key will be generated but the data will not be uploaded to avoid de-duplication(data duplication).Then the TPA checks the data authentication also provide security to data. As the data is uploaded on the server load balancing take place at each server.

VISUAL CRYPTOGRAPHY:

Visual Cryptography is a cryptographic technique where visual Information (Image, text, etc) gets encrypted in such a way that the decryption can be performed by the human visual system without aid of computers. It is also used to protect image-based secret information.The technique random sequence is used to divide the sequence to divide an image into n number of shares. The shares are sent through different communication channels from sender to receiver so that the probability of getting sufficient shares by the intruder minimizes.

Advantages of Visual Cryptography:

1. Encryption doesn't required any NP hard problem dependency.
2. Decryption algorithm not required. So the unknown person unknown to cryptography can decrypt the message.
3. We can send cipher text through FAX or E-MAIL.
4. Infinite Computation Power can't predict the message.

VII ALGORITHM:

1. RSA:

RSA is one of the first practical public-key cryptosystems and is widely used for secure data transmission. In such a cryptosystem, the encryption key is public and differs from the decryption key which is kept secret. In RSA, this asymmetry is based on the practical difficulty of factoring the product of two large prime numbers, the factoring problem. RSA is made of the initial letters of the surnames of Ron Rivest, Adi Shamir, and Leonard Adleman, who first publicly described the algorithm in 1977. Clifford Cocks, an English mathematician, had developed an equivalent system in 1973, but it was not declassified until 1997. A user of RSA creates and then publishes a public key based on two large prime numbers, along with an auxiliary value. The prime numbers 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 numbers can feasibly decode the message. Breaking RSA encryption is known as the RSA problem; whether it is as hard as the factoring problem remains an open question

2. K-N Sharing:

In this algorithm the cryptography created by Adi Shamir. It is a form of secret sharing, where a secret is divided into parts, giving each participant its own unique part, where some of the parts or all of them are needed in order to reconstruct the secret.Counting on all participants to combine together the secret might be impractical, and therefore sometimes the threshold scheme is used where any k of the parts are sufficient to reconstruct the original secret.

3.SHA-1

SHA-1 produces a 160-bit (20-byte) hash value known as a message digest. A SHA-1 hash value is typically rendered as a hexadecimal number, 40 digits long.

SHA-1 is a member of the Secure Hash Algorithm family. The four SHA algorithms are structured differently and are named SHA-0, SHA-1, SHA-2, and SHA-3. SHA-0 is the original version of the 160-bit hash function published in 1993 under the name SHA: it was not adopted by many applications. Published in 1995, SHA-1 is very similar to SHA-0, but alters the original SHA hash specification to

ISSN: 2278 – 1323

correct weaknesses that were unknown to the public at that time. SHA-2, published in 2001, is significantly different from the SHA-1 hash function.

4.Load balancing(computing): In computing, **load balancing** distributes workloads across multiple computing Using multiple components with load balancing instead of a single component may increase reliability and availability through redundancy. Load balancing usually involves dedicated software or hardware, such as a multilayer switch or a Domain Name System server process.Load balancing differs from channel bonding in that load balancing divides traffic between network interfaces on a network socket (OSI model layer 4) basis, while channel bonding implies a division of traffic between physical interfaces at a lower level, either per packet (OSI model Layer 3) or on a data link (OSI model Layer 2) basis with a protocol like shortest path bridging.

5. Huffman coding:

In computer science and information theory, a **Huffman code** is a particular type of optimal prefix code that is commonly used for lossless data compression. The process of finding and/or using such a code proceeds by means of **Huffman coding**, an algorithm developed by David A. Huffman while he was a Ph.D. student at MIT, and published in the 1952 paper "A Method for the Construction of Minimum-Redundancy Codes".

The output from Huffman's algorithm can be viewed as a variable-length code table for encoding a source symbol (such as a character in a file). The algorithm derives this table from the estimated probability or frequency of occurrence (*weight*) for each possible value of the source symbol. As in other entropy encoding methods, more common symbols are generally represented using fewer bits than less common symbols. Huffman's method can be efficiently implemented, finding a code in linear time to the number of input weights if these weights are sorted.[2] However, although optimal among methods encoding symbols separately, Huffman coding is not always optimal among all compression methods.

6.MD5:MD5 message-digest algorithm is a widely used cryptographic hash function producing a 128-bit (16-byte)hash value, typically expressed in text format as a 32 digit hexadecimal number. MD5 has been utilized in a wide variety of cryptographic applications, and is also commonly used to verify data integrity.

MD5 was designed by Ronald Rivest in 1991 to replace an earlier hash function, MD4 The source code in RFC 1321 contains a "by attribution" RSA license.

In 1996 a flaw was found in the design of MD5. While it was not deemed a fatal weakness at the time, cryptographers began recommending the use of other algorithms, such as SHA-1—which has since been found to be vulnerable as well.[4] In 2004 it was shown that MD5 is not collision resistant. As such, MD5 is not suitable for applications like SSL certificates or digital signatures that rely on this property for digital security. Also in 2004 more serious flaws were discovered in MD5, making further use of the algorithm for security purposes questionable; specifically, a group of researchers described how to create a pair of files that share the same MD5checksum. Further advances were made in breaking MD5 in 2005, 2006, and 2007.

CONCLUSION

In this project we will propose a system that will eliminate issues like de-duplication, data authentication and security using visual cryptography.

REFERENCES

1. Q. Zheng and S. Xu, "Secure and Efficient Proof of Storage with De duplication," in Proc. ACM Conference on Data and Application Security and Privacy (CODASPY), 2012.

2 B. Chen, R. Curtmola, G. Ateniese, and R. Burns, "Remote Data Checking for Network Coding-based Distributed storage Systems," in Proc. ACM Cloud Computing Security Workshop (CCSW),2010, pp. 31–42.

3.N. Cao, S. Yu, Z. Yang, W. Lou, and Y. T. Hou, "LT Codes based Secure and Reliable Cloud Storage Service," in Proc. IEEE International Conference on Computer Communications (INFOCOM),2012.

4. S. D. C. di Vimercati, S. Foresti, S. Paraboschi, G. Pelosi, and P. Samarati, "Efficient and Private Access to Outsourced Data," in Proc. IEEE International Conference on Distributed Computing Systems (ICDCS), 2011, pp. 710–719.

5. Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S.Yau, "Dynami Audit Services for Integrity Verification of Outsourced Storage in Clouds," in Proc. ACM Symposium on Applied Computing (SAC), 2011, pp. 1550–1557.

6. S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Secure, Scalable, and Fine-grained Data Access Control in Cloud Computing," in Proc. IEEE International Conference on Computer Communications INFOCOM), 2010, pp. 534–542.

ISSN: 2278 – 1323