ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

# Survey on Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage

Tessy Vincent, Chippy Jacob, Rakhi Ramachandran Nair, Krishnaveni.V.V

M.Tech Student, Dept. of Computer Science and Engineering, College of Engineering, Kidangoor, Kottayam, Kerala, India

M.Tech Student, Dept. of Computer Science and Engineering, College of Engineering Kidangoor Kottayam, Kerala, India

M.Tech Student, Dept. of Computer Science and Engineering, College of Engineering, Kidangoor Kottayam, Kerala, India

Assistant Professor, Dept. of Information Technology, College of Engineering, Kidangoor Kottayam, Kerala, India

**ABSTRACT:** the protection of outsourced data in cloud storage from corruptions is becomes critical. Fault tolerance must be provided to cloud storage together with data integrity checking and failure reparation. Recently while providing fault tolerance, due to lower repair bandwidth, regenerating codes have gained popularity. Today the existing remote checking methods for regenerating-coded data only provide private auditing. That is, data owners are required to always stay online for auditing and failure repairing, which is sometimes difficult and impractical. Here, a public auditing scheme for the regenerating code based cloud storage is proposed. In the proposed system, a proxy is introduced into the traditional auditing system models which have the privilege to regenerate the authenticators, to solve the regeneration problem of failed authenticators. Apart from this, a public verifiable authenticator is presented, which is generated by a pair of keys and can be regenerated using partial keys, to help data owners to get released from online burden. Also to randomize the encode coefficients for data privacy, a pseudorandom function is used. Extensive security analysis shows that this scheme is provable secure under random oracle model and the experimental evaluation indicates that this scheme is highly efficient and feasible to integrate into regenerating-code-based cloud storage.

**KEYWORDS***:* Cloud storage, regenerating codes, public audit, privacy preserving, authenticator regeneration, proxy, privileged, Provable secure.

## I. INTRODUCTION

In recent years, cloud storage service has become a faster profit growth providing a comparably low cost, scalable, position independent platform for clients' data. This new paradigm of data hosting brings some security threats towards users' data, thus making feel difficulty. Thus, the correctness, availability and integrity of the data are at risk. On the one hand, the cloud service is facing a broad range of internal/external adversaries, who would delete or corrupt users' data; on the other hand, the cloud service providers may act dishonestly, trying to hide data loss or corruption and claiming that the files are still correctly stored in the cloud for monetary and reputation reasons. Thus it is needed for the users to implement an efficient protocol to perform periodical verifications of their outsourced data to ensure the correctness of data stored in cloud.

Many techniques dealing with the integrity of data on cloud without a local copy have been proposed under different system and security models. The most promising work among these studies are the PDP (*provable data possession*) model and POR (*proof of retrievability*) model, which were actually proposed for the single-server scenario by Ateniese *et al.* [1] and Juels *et. al.* [2].

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

Considering that files are usually striped and redundantly kept across multi-clouds or multi-servers, *R.Curtmola et.al* [3]explore integrity verification schemes appropriate for such multi-servers or multiclouds setting with various redundancy schemes, like *replication*, *erasure codes*, and, more recently, *regenerating codes*. *Jian Liu* et.al [9] mainly concentrates on the integrity verification problem in regenerating code based cloud storage especially with the functional repair strategy. Several other studies extended the single-server scheme to the regenerating-code-scenario and designed and implemented data integrity protection scheme cloud storage and the scheme is changed to the thin cloud setting. However, most methods are designed for private audit, allowing only the data owner to verify the integrity and repair the faulty servers. The tasks of auditing and reparation in the cloud can be difficult and expensive for the users by considering the large sized outsourced data and the user's restricted resource capability. It should minimize the overhead of using cloud storage as much as possible so that a user needs not to perform too many operations to the outsourced data. Particularly, users do not want to go through the complexity in verifying and reparation. The auditing schemes imply the problem that users need to stay online, which may impede its implementation, especially for long-term archival or depository storage.

To completely ensure the data integrity and save the users' computational resources and online burden, here a public auditing scheme for the regenerating code based cloud storage, in which the integrity checking and regeneration are implemented on behalf of the data owner by a third party auditor and a semi trusted proxy separately. Instead of directly adapting the existing public auditing scheme to the multi-server setting, another novel authenticator is designed, which is more suitable for regenerating codes. Besides ,encrypt the coefficients to protect data privacy against the third party auditor, which is more simple than applying the proof blind technique and data blind method in [7].Several challenges and threats suddenly arise in the new system model with a proxy but security analysis shows that this system works well with these problems.

## II. LITERATURE SURVEY

*G. Ateniese et al.* introduces a model for provable data possession (PDP) that provides probabilistic proof that a third party stores a file. Also presents a Sampling Provable Data Possession (SPDP) scheme, which combines the RSA cryptography with Homomorphic Verifiable Tags (HVT). This model gives the provably secure scheme for remote data checking. A PDP protocol checks that an outsourced storage site retains a file, which consists of a collection of blocks. The client (data owner) pre-processes the file, generating a piece of metadata that is stored locally, transmits the file to the server and may delete its local copy. Before deleting its local copy of the file, the client may execute a data possession challenge to make sure the server has successfully stored the file. Clients may encrypt a file prior to outsourcing the storage. For this purpose, encryption is an orthogonal issue; the file may consist of encrypted data and metadata does not include encryption keys. At a later time, the client issues a challenge to the server to establish that the server has retained the file. The client requests that the server compute a function of the stored file, which it sends back to the client. Using its local metadata, the client verifies the response.

The goal of a PDP scheme is to detect server misbehavior when the server has deleted a fraction of the file. The performance of PDP is bounded by disk I/O and not by cryptographic computation. Although the SPDP scheme can keep the data privacy, it cannot support the dynamic auditing and the batch auditing for multiple owners.

*A. Juels et al.* "PORs: Proofs of Retrievability for Large Files" presents POR is a protocol in which the verifier stores only a single cryptographic key irrespective of the size and number of the files. This scheme requires that the prover access only a small portion of a (large) file F in the course of a POR. This POR protocol encrypts F and randomly embeds a set of randomly-valued check blocks called *sentinels*. The use of encryption here renders the sentinels indistinguishable from other file blocks. The verifier challenges the prover by specifying the positions of a collection of sentinels and asking the prover to return the associated sentinel values. If the prover has modified or deleted a substantial portion of F, then with high probability it will also have suppressed a number of sentinels. It is therefore unlikely to respond correctly to the verifier. To protect against corruption by the prover of a small portion of F, error-correcting codes are also employed. A drawback of POR scheme is the preprocessing / encoding of F required prior to storage with the prover. This step imposes some computational overhead beyond that of simple encryption or hashing as well as larger storage requirements on the prover. The sentinels may constitute a small fraction of the encoded F' (typically, say, 2%); the error-coding imposes the bulk of the storage overhead.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

*R. Curtmola et al.* "MR-PDP: Multiple Replica Provable Data Possession", represents a storage system replication that increases durability and availability of the data. This provides no stronger evidence that multiple copies of the data are actually stored. It presents MR-PDP techniques that present t replicas to verify the challenges. This unique replica can produce at t unique time and that system uses t time to store the t unique replica. It shows that t replica is much more efficient than using single replica.

*Kevin D.Bowers et al.* "HAIL: A High- Availability and Integrity Layer for Cloud Storage" introduce the High Availability and Integrity Layer (HAIL) for distributed cryptographic system that allows a set of server that prove to client that file stored is retrievable and intact. HAIL verifies and reactively reallocates file shares cryptographically. This paper shows how HAIL enhances the security and efficiency of existing tools, like Proofs of Retrievability (POR) implemented on individual servers.

*K. Yang et al.* "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing": In cloud computing data owner host their data on cloud server and user can access their data from cloud server. Due to this new paradigm of outsourcing, the new security challenges which require checking the data integrity in cloud storage increases. Some techniques are available for static data storage but for dynamic data storage an efficient and secure dynamic auditing protocol is desired to convince data owners that the data are correctly stored in the cloud. For that here designed an auditing framework for cloud storage system and presents an efficient privacy preserving auditing protocol. This auditing protocol to support the dynamic data operations, which is efficient and provably secure in the random oracle model. Also this auditing protocol support for both multiple owners and multiple clouds, without using any trusted organizer. This paper implements secure dynamic auditing protocol.

*Y. Zhu et al.* "Cooperative Provable Data Possession for Integrity Verification in Multicloud Storage" presents a cooperative PDP (CPDP) scheme based on homomorphic verifiable response and hash index hierarchy method. Also verify the security of the scheme based on multiprover zero-knowledge proof system, which can satisfy completeness, knowledge soundness and zero knowledge properties. In addition, performance optimization mechanisms are intellectual for the scheme, and in particular present an efficient method for selecting optimal parameter values to minimize the computation costs of clients and storage service providers.

*C. Wang et al.* "Toward Secure and Dependable Storage Services in Cloud Computing", proposed a flexible distributed storage integrity auditing mechanism, utilizing the homomorphic token and distributed erasure coded data. This scheme allows users to examine the cloud storage with very lightweight communication and computation cost. The auditing result not only ensures correctness guarantee of cloud storage, but also at the same time achieves fast data error localization, i.e., the identification of misbehaving server. Considering the dynamic nature of cloud data, this system design also supports secure and efficient dynamic operations on outsourced data, including block deletion, modification and append. It is based on the observation of linear property of the parity vector blinding process. Recall that the reason of blinding process is for protection of the secret matrix against cloud servers. However, this can be achieved either by blinding the parity vector or by blinding the data vector (assume $k < m$). Thus, if data vector is made blind before file distribution encoding, then the storage verification task can be successfully delegated to third party auditing in a privacy preserving manner. Specifically, the new protocol is described as follows:

1. Before file distribution, the user blinds each file block data.
2. Based on the blinded data vector the user generates k parity vectors via the secret matrix.
3. The user calculates the i'th token for server j.
4. The user sends the token set, secret matrix P, permutation and challenge key, to TPA for auditing delegation.

TPA does not have to take away the blinding values in the servers' response but verifies directly. As TPA does not know the secret blinding key there is no way for TPA to learn the data content information during auditing process. Therefore, the privacy-preserving third party auditing is achieved, but the overall computation overhead and communication overhead remains roughly the same.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

*S. G. Worku et al.* "Secure And Efficient Privacy-Preserving Public Auditing Scheme for Cloud Storage" presents a privacy preserving public auditing scheme that supports public auditing and identity privacy on shared data stored in the cloud storage service for enhancing its security and effectiveness. Using HARS and its properties, designs a privacy preserving public auditing technique for outsourced data in the cloud. Here, the TPA can verify the completeness of shared data for a group of users and the identity of the signer on each block in shared data is kept private from the TPA during the auditing.

*Ring Signatures:* The concept of ring signatures is that with this, a verifier agreed that a signature is computed using one of group members' private keys, but the verifier is not able to find which one. This property can be used to keep the identity of the signer from a verifier.

*Homomorphic Authenticable Ring Signature:* The ring signatures generated by HARS is able not only to preserve identity privacy but also to support blockless verification. HARS consist of three algorithms: KeyGen, RingSign and RingVerify. In KeyGen, each user in the group generates her public key and private key. In RingSign, a group user can sign a block with her private key and all the group members' public keys. A verifier is allowed to identify whether a given block is signed by a group member in RingVerify.

*Generation of Key and Signature for user file:* In Key-Gen, users generate their own public/private key pairs. In SigGen, a user is able to find ring signatures on blocks in shared data. Each user in the group can perform an insert, delete or update operation on a block, and determine the new ring signature on this new block in Modify. MD5 operation is employed for key generation. To generate a ring signature, find out the symmetric key, Pick a random initial value, Pick random xi's and solve for ys and the signer's trapdoor permutation is inverted. Here AES algorithm is used for file encryption and uploaded to the cloud and TPA.

*Jian Liu et al.* "Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage" presents the auditing system model for Regenerating Code based cloud storage involves four entities as in Fig: 1

1. *the data owner*, a person who owns or stores large amounts of data or files that to be stored in the cloud;
2. *the cloud*, provide storage service for the users to store cloud data and have significant computational resources that are managed by the cloud service provider.
3. *The third party auditor* (TPA), who has the rights and capabilities to conduct public audits on the coded data in the cloud. The TPA is trusted, reliable and independent. TPA should check the data integrity and availability at frequent time intervals.TPA should be allowed for organizing, managing, and maintaining the shared data instead of data owners. It also makes sure that it does not disturb data owners and its audit result is unbiased for both data owners and cloud servers.
4. *a proxy agent*, who is semi-trusted and acts on behalf of the data owner on the failed servers during the repair procedure to regenerate authenticators and data blocks. The data owner is limited in computational and storage resources as compared to other entities and may become off-line even after the data upload procedure. The proxy, would be always on-line, is meant to be rather more powerful than the data owner however less than the cloud servers in terms of computation and memory capacity. To save resources and the online burden potentially caused by the periodic auditing and unexpected repairing, the information owners ask to the TPA for integrity verification and delegate the reparation to the proxy.

Auditing scheme:

The proposed auditing scheme mainly consists of three procedures: Setup, Audit and Repair. Each procedure contains some polynomial time algorithms as follows:

**Setup:** The data owner keeps this procedure to initialize the auditing scheme related parameters.

*KeyGen (1k) → (pk, SK)*: This polynomial-time algorithm is run by the data owner to initialize its public and secret parameters by taking a security parameter $\kappa$ as input.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

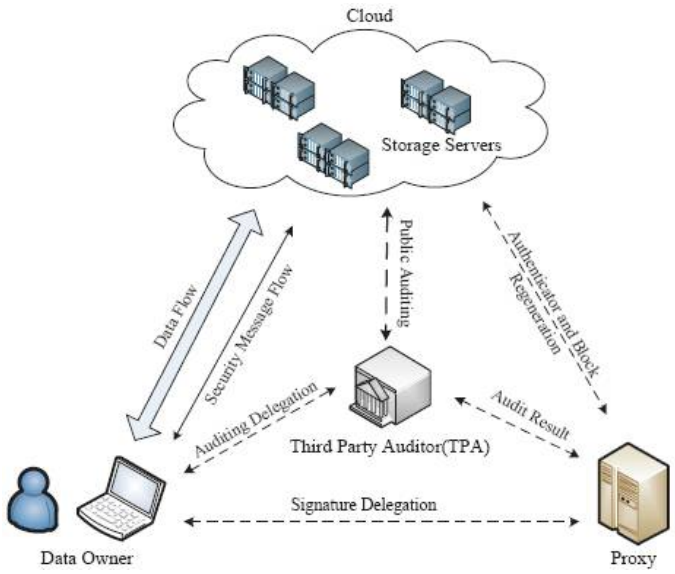


Fig.1 The system model

*Degelation (SK) → (x)*: This algorithm represents the interaction between the data owner and proxy. The data owner generates partial secret key *x* to the proxy through a secure approach.

*SigAndBlockGen(sk,F) → (ϕ, Ψ, t):* This polynomial time algorithm is run by the data owner and takes the secret parameter sk and the original file F as input, and then outputs a coded block set Ψ, an authenticator set ϕ and a file tag t.

**Audit:** The cloud servers and TPA interact with one another to take a random sample on the blocks and check the data intactness in this procedure.

*Challenge (Finfo) → (C)*: This algorithm is performed by the TPA with the information of the file *Finfo* as input and a challenge *C* as output.

*ProofGen(C, ϕ, Ψ) → (P)*: This algorithm is run by each cloud server with input challenge *C*, coded block set Ψ and authenticator set ϕ, then it outputs a proof *P*.

*Verify (P, pk, C) → (0, 1)*: This algorithm is run by TPA immediately after a proof is received. Taking the proof *P*, public parameter *pk* and the corresponding challenge *C* as input, it outputs 1 if the verification successfully passed and 0 otherwise.

**Repair:** In the absence of the data owner, the proxy interacts with the cloud servers during this procedure to repair the wrong server detected by the auditing process.

*ClaimForRep (Finfo) → (Cr):* This algorithm is similar with the Challenge () algorithm in the Audit phase, but outputs a claim for repair Cr.

*GenForRep (Cr, ϕ, Ψ) → (BA):* The cloud servers run this algorithm upon receiving the Cr and finally output the block and authenticators set BA with another two inputs ϕ, Ψ.

*BlockAndSigReGen(Cr, BA) → (ϕ' ,Ψ',⊥):* The proxy implements this algorithm with the claim Cr and responses BA from each server as input, and outputs a new coded block set and authenticator set if successful, outputting ⊥ if otherwise.

## III. CONCLUSION

In this survey, different papers regarding privacy preserving public auditing for regenerating code based cloud storage is studied. In this public auditing scheme for the regenerating code based cloud storage system, the data owners have the right to delegate TPA for checking their data validity. To protect the original data privacy against the TPA, the coefficients are randomized in the beginning rather than applying the blind technique during the auditing process.

ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization Volume 5, Special Issue 14, December 2016

3rd National Conference on Recent Trends in Computer Science and Engineering and Sustainability in Civil Engineering (TECHSYNOD'16)

19th, 20th and 21st December 2016

Organized by

Department of CSE, MCA &CE, Mohandas College of Engineering and Technology, Thiruvananthapuram, Kerala, India

## REFERENCES

[1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 598-609.

[2] A. Juels and B. S. Kaliski Jr, "Pors: Proofs of retrievability for large files," in *Proceedings of the 14th ACM conference on Computer and communications security*. ACM, 2007, pp. 584–597.

[3] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "Mr-pdp: Multiple replica provable data possession," in *Distributed Computing Systems, 2008. ICDCS'08. The 28th International Conference on*. IEEE, 2008, pp. 411–420.

[4] K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availability and integrity layer for cloud storage," in *Proceedings of the 16th ACM conference on Computer and communications security*. ACM, 2009, pp. 187–198.

[5] K. Yang and X. Jia, "An efficient and secure dynamic auditing protocol for data storage in cloud computing," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 24, no. 9, pp. 1717–1726, 2013.

[6] Y. Zhu, H. Hu, G.-J. Ahn, and M. Yu, "Cooperative provable data possession for integrity verification in multicloud storage," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 23, no. 12, pp. 2231–2244, 2012.

[7] C. Wang, Q. Wang, K. Ren, and W. Lou, "Towards secure and dependable storage services in cloud computing," *Service Computing, IEEE Transactions on*, vol. 5, no. 2, pp. 220–232, May 2012.

[8] S. G. Worku, C. Xu, J. Zhao, and X. He, "Secure and efficient privacy preserving public auditing scheme for cloud storage," *Computers & Electrical Engineering*, 2013.

[9] Jian Liu, Kun Huang, Hong Rong, Huimei Wang And Ming Xian," Privacy-Preserving Public Auditing for Regenerating-Code-Based Cloud Storage", I*nformation And Security, IEEE Transactions on*, vol 1 no 2015.