Privacy Preserving Dynamic Provable Data Possession with Batch Update for Secure Cloud Storage

Smita Chaudhari1*, Gandharba Swain2
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur, Andhra Pradesh, India

Abstract—Cloud Server (CS) is an untrusted entity in cloud paradigm that may hide accidental data loss to maintain its reputation. Provable Data Possession (PDP) is a model that allows Third Party Auditor (TPA) to verify the integrity of outsourced data on behalf of cloud user without downloading the data files. But this public auditing model faces many security and performance issues such as: unnecessary computational burden on user as well as on TPA, to preserve identities of users from TPA during auditing, support for dynamic updates etc. Many PDP schemes creates computational burden either on TPA or Cloud User. To balance this overhead between TPA and User, this paper proposes Privacy-Preserving Dynamic Provable Data Possession (P2DPDP) scheme, which is based on ODPDP scheme. In ODPDP scheme, user relieves the burden by signing a contract with TPA regarding verification of his outsourced data. But this scheme generates computation overhead on TPA. To reduce this computation overhead of TPA, our P2DPDP scheme uses Indistinguishability Obfuscation (IO) with one-way function such as message authentication code to make a lightweight auditing process. P2DPDP scheme uses Rank-based Merkle Tree (RBMT) to support dynamic updates in batch mode which greatly reduces computation overhead of TPA. ODPDP lacks privacy which is maintained in P2DPDP using ring signature technique. Our experimental results demonstrate the reduced verification time and computation cost compared to existing schemes.

Keywords—Public auditing; ring signature; Indistinguishability Obfuscation; Rank-based Merkle Tree (RBMT)

I. INTRODUCTION

With the rapid development of cloud computing, many individuals or small-scale organizations started outsourcing their data on untrusted CS. This paradigm even though, proved to be a boon, has brought many security challenges with it. The user is not having physical ownership of data since outsourced data may be stored at any server. This outsourced data may get damaged unintentionally because of disk crashes or natural disasters. Sometimes CS may delete infrequent data intentionally to create space for new users. These incidents or data loss are not reported to cloud users to maintain reputation. Provable Data Possession (PDP) is a technique that allows any user to check the integrity of data blocks outsourced on CS without downloading the entire data file. Many researchers have proposed cryptographic techniques using homomorphic authenticators [2]-[12]. Cloud users can check the integrity of outsourced data on their own but this frequent task creates an additional burden on the cloud user in terms of time and computation cost. So, researchers have proposed solution in which user can delegate auditing work to TPA having expertise and skill. TPA generates a challenge message and CS has to produce the proof based on recent data. This proof is verified by TPA. This model has many security challenges since CS is one of the untrusted entities and TPA even if trusted, curious about auditing work.

Integrity checking during dynamic update operations of data is a major challenge in the PDP model. Authenticators have to be recalculated during insertion, deletion, and modification operations on data because most of the authenticators are calculated based on indices of files. Researchers have proposed dynamic auditing schemes using Merkle Hash Tree (MHT) [24]. MHT is a hash-based data structure used for data verification. Another data structure used for dynamic data updates are Index Hash Table (IHT) [5],[6] Dynamic Hash Table (DHT) [18]. These techniques need some additional information to store which may increase storage and computation cost. Guo et al. [21] proposed Multi-leaf-authenticated (MLA) scheme for dynamic data using Rank based Merkle Tree (RBMT). This scheme authenticates multiple leaf nodes at once without storing height and status value. With this scheme, multiple dynamic update operations can be performed in batch mode.

In auditing model, TPA is one of the trusted entity but still curious and may compromise data and user privacy. During auditing, TPA may infer user information who have signed the blocks. Researchers have given multiple approaches to address this issue. Some privacy-preserving protocols [39],[40] are proposed based on aggregate signature [14] or hash-based commitment [15] but auditing is not mentioned in these schemes. Huang et al. [18] proposed privacy-preserving scheme using group signature and blockchain. Tian et al. [20] also proposed a privacy-preserving scheme that addresses data privacy using random masking to blind the data proof. Identity privacy is preserved using a modification record table to record operation information. Different variations of signature algorithms, such as blind signature, random sampling, ID-based privacy, and attribute-based signature are used by multiple researchers during auditing [22]-[28]. Attribute-based signatures are based on attribute-based cryptography [16],[17]. Ring signature is another variation of group signature in which ring or group is formed with multiple users. User sign the blocks and share it among the group members. Verifier determines that block is signed by one of the group members.

*Corresponding Author
but can’t reveal who has signed the block. Certificateless authentication [29],[30] is one of the cryptographic techniques for authentication. Some authors proposed privacy-preserving integrity verification scheme [31],[32] using certificate-less ring signature which greatly reduces the computation cost during auditing. Li et al. [19] proposed integrity checking of group shared data using certificateless signature. Ni et al. [33] proposed lightweight ID-P^DP scheme in which privacy is achieved through zero-knowledge proof.

Many of the above auditing schemes are based on homomorphic authenticators which incur high computation cost and time during auditing. There is a need to propose a lightweight auditing process which can create a little burden on TPA as well as on cloud user. Indistinguishability Ofuscation (IO) is a modern cryptographic technique that uses one-way function for implementation of different cryptography constructs [35]. Researchers [36]-[38] proposed an efficient public verification scheme using IO combined with one-way function MAC which makes this scheme lightweight by generating very little burden on TPA. Tu et al. [13] proposed user-focus auditing which try to reduce the overhead of user by pre-generating challenges for TPA before auditing. Guo et al. [21] also proposed outsourced auditing scheme in which after each verification, TPA generates an audit data log which is checked later by the user. These schemes reduce the burden of verification on user. There is no need for a user to be available during verification, as per his convenience he can check the audit log.

Most of the auditing schemes create computational burden of tag generation and verification either on TPA or user. If we try to reduce the burden of TPA, it will increase the burden on user or vice-versa. Zhang et al. [36] scheme reduces the computation overhead of TPA but cloud user is actively involved in auditing process. Guo et al. [21] scheme proposes auditing scheme where user is not involved in auditing process but it creates computation burden on TPA because of Efficient Homomorphic Verifiable Tags (EHVT).

The remainder of this paper is described as follows: Section II describes related work, problem identification and contributions by authors. Section III elaborates basic building blocks of P^DPDP scheme. Section IV explains the P^DPDP scheme. Section V discusses the evaluation of P^DPDP scheme in terms of security and performance.

II. RELATED WORK, PROBLEM IDENTIFICATION AND RESEARCH CONTRIBUTION

A. Related Work

A good number of solutions have been proposed by many researchers [1]-[10] for integrity verification of outsourced data. In most of these schemes, cloud user and TPA are actively involved during verification phase. This may create an additional burden on cloud users as well as on TPA. Guo et al. [21] proposed ODPDP scheme which relieves user’s verification overhead by migrating frequent auditing tasks to TPA. In this scheme, a contract takes place between user and TPA regarding the frequency of verification task. TPA generates challenges based on this contract. After each verification, a log file is generated at TPA which contains an audit data log. User as per his convenience can check the log and make sure the integrity of his data as well as working of TPA. This scheme greatly reduces the overhead on user side. This scheme uses Multi-Leaf Authentication (MLA) solution with RBMT for dynamic data updates which greatly reduces storage cost as well as allows verification of multiple dynamic operations in batch mode.

To minimize the burden of TPA in terms of computations, it is necessary to develop a verification scheme which is lightweight in terms of computation. Zhang et al. [36] proposed lightweight auditing technique using IO and one-way function MAC. This greatly reduces the computation overhead of TPA during verification since TPA has to just calculate MAC each time and verify it with the received proof from CS. In this scheme, during outsourcing data at CS, a user has an additional overhead of generating circuit (audit program), obfuscating with MAC key, and send it to CS. But this is only a one-time cost to generate obfuscated program since it would not change along with the modification of public parameters and challenge message. This scheme uses Merkle Hash Tree (MHT) to support dynamic data updates on file.

In auditing model, generating privacy-preserving auditing technique is a major challenge because of the curious but trusted nature of TPA. Thokcham [34] proposed a privacy-preserving auditing technique using CDH based ring signature. This ring signature scheme is unforgeable and completely anonymous. Any ring member can sign a message using his private key and public keys of other members. So not every member needs to be present during the signing process. Using this scheme, anyone can check whether a signature is generated or not without revealing the user’s identity who has signed that message. Thus, preserving the identity privacy of user during verification from TPA.

ODPDP scheme reduces user overhead but increases burden on TPA in terms of computation. The computation cost on TPA in ODPDP during auditing is (l+s+1)ExpG + 2Pair. Where l is number of challenged blocks, s denotes number of sectors per block, ExpG is exponentiation over group G and Pair is pairing operation. Compare to ODPDP scheme, Zhang et al. [36] scheme create less burden on TPA during auditing i.e., 2 Hashzp where Hash is hashing operation on Zp. It means TPA has to compute only 2 hash functions for verification. So proposed P^DPDP scheme modifies the ODPDP by using IO and MAC proposed by Zhang et al. scheme instead of EHVT of ODPDP for integrity verification. ODPDP scheme not proposing any solution for identity privacy. We extend this scheme by using CDH based ring signature to achieve identity privacy proposed by Thokcham [34]. P^DPDP also uses RBMT of ODPDP which allows verification of multiple dynamic operations in batch mode compared to MHT in Zhang et al. [36].

B. Problem Identification

To balance the computation overhead between user and TPA, this paper proposes Privacy-Preserving Dynamic Provable Data Possession (P^DPDP) scheme for cloud storage. In this scheme, the main purpose of using IO is to reduce the computation burden of TPA while maintaining security. Since
TPA only needs to validate the commitments generated by CS, user’s data will not be revealed to TPA which preserve data privacy during auditing process. To avoid the continuous involvement of user during auditing process, Cloud user and TPA sign a contract which includes starting address of the block, the frequency at which auditor launches a challenge, and number of challenged blocks. TPA verifies the outsourced data based on contract and generates a log file which can be verified by user as per his convenience. For the support of dynamic updates, RBMT is used that can perform multiple update operations in a batch way. P^2DPDP support user groups and preserve identity privacy by using CDH based ring signature scheme.

C. Research Contribution

Specifically, the contribution of our scheme P^2DPDP is as follows:

- Guo et al. [21] proposed ODPDP scheme which uses Effective Homomorphic Verifiable Tag (EHVT) for integrity verification and RBMT for dynamic data processing. To make the auditing process lightweight and to reduce computation overhead of TPA, we modify the integrity verification scheme of ODPDP by using indistinguishability obfuscation of Zhang et al. [36].

- Guo et al. [21] scheme not offering a solution for identity privacy during auditing. We extend this ODPDP scheme to achieve group user privacy using CDH based ring signature.

- We describe a concrete P^2DPDP scheme to be secure and lightweight by modifying ODPDP scheme. Experimental results certify the performance of our scheme.

III. PRELIMINARIES

This section introduces cryptographic building blocks used in P^2DPDP scheme such as IO for integrity verification, MLA for dynamic updates and CDH based ring signature scheme for privacy-preserving.

A. Indistinguishability Obfuscation (IO)

Indistinguishability obfuscation is a notion that obfuscates any two distinct (equal size) programs that implement identical functionalities but computationally indistinguishable from each other [35].

Assume \(\{C_i\}\) is a circuit class with security parameters \(l\). A uniform PPT algorithm \(iO\) having input \(l\), circuit \(C \in \{C_i\}\) and outputs a circuit \(C’\) is called indistinguishable obfuscator if the following conditions are fulfilled:

1) For all security parameters \(l\), Circuit \(C\), and input \(x\), we have probability as:
\[
\Pr[C’(x)=C(x)]=1, \text{ where } C’=iO(l,C) \tag{1}
\]

Equation (1) satisfies the completeness property of IO. It states that circuit \(C’\) must behave exactly same as circuit \(C\) if \(C’\) is generated by an independent invocation of \(iO\) on \(C\).

2) For any (not essentially uniform) PPT adversaries \(D\), for all security parameter \(l \in \mathbb{N}\), for all pairs of circuits \(C_0, C_1 \in C_1\), there exists a negligible function \(Negl()\) such that if \(C_0(x)=C_1(x)\) for all inputs \(x\), then:
\[
\Pr[D(iO(l, C_0))=1] - \Pr[D(iO(l, C_1))=1] \leq Negl(l) \tag{2}
\]

Equation (2) satisfies the indistinguishability property of IO. It states that the secrets embedded in obfuscated program cannot be extracted by \(D\).

Zhang et al. [36] proposed integrity verification using IO and one-way function MAC.

B. Multi-Leaf Authentication (MLA)

ODPDP scheme [21] proposed MLA for dynamic updates. This scheme uses RBMT instead of MHT to support authentication of indices of leaf nodes. In RBMT, no need to store height value as in MHT. Each node contains only two fields \((r,h)\) where \(r\) is the rank of a node which is the number of leaf nodes reachable from node \(\omega\) and \(h\) is a hash value of that node. Mainly, the rank of a leaf node is 1 i.e., \(r=1\). The second element \(h\) is defined as in (3):
\[
h = \left\{ \begin{array}{ll}
H_2(m_i), & \text{if } \omega \text{ is the } i\text{th leaf node} \\
H_1(r||\omega.\text{left.hash}||\omega.\text{right.hash})), & \text{if } \omega \text{ is a non-leaf node} \tag{3}
\end{array} \right.
\]

where \(H_1\) and \(H_2\) are secure collision-resistant hash function and || denotes concatenation. The outsourced file \(F\) is divided into \(n\) blocks such as \(F = \{f_1, f_2, \ldots, f_n\}\). The \(i\)th element \(f_i\) is bind to the \(i\)th leaf node of RBMT by storing the hash value of \(f_i\) at node \(\omega_i\) using \(H_2\) in (3). Thus, leaf nodes are already sorted from left to right by their indices. For each non-leaf node, \(\omega.\text{left.hash}\) and \(\omega.\text{right.hash}\) indicates the hash value of the left node and right node respectively calculated using \(H_1\) in (3). RBMT can be constructed for given \(n\) data blocks. A Merkle root \(h_{\text{root}}\) is sufficient to check the integrity of dynamic updates in a tree because of the dependency of all data blocks. Fig. 1 shows the RBMT constructed over 14 data blocks. In Fig.1, when multiple leaf nodes are challenged using MHT such as \(\omega_3\) and \(\omega_7\), the proofs generated are \(\Omega_1 = \{\omega_{26}, \omega_{22}, \omega_{15}, \omega_{24}, \omega_3\}\) and \(\Omega_2 = \{\omega_{26}, \omega_{21}, \omega_{17}, \omega_{18}, \omega_1\}\). During verification using MHT, some repeated and unnecessary nodes have to be retrieved and processed that incur large computation costs. But using MLA solution, if multiple challenged nodes are \((3,7,8,10,13)\), the corresponding multi-proof \(L_p\) is:
\[
L_p = \{\omega_3, \omega_7, \omega_8, \omega_{10}, \omega_{13}, \omega_{16}, \omega_{18}, \omega_{19}, \omega_{21}, \omega_{22}, \omega_{23}, \omega_{24}, \omega_{25}, \omega_{26}\}
\]

where every necessary node appears just once which reduces computation cost as well as support multiple updates in batch mode.

C. CDH based Ring Signature Scheme

To achieve privacy during auditing, Thokcham [34] used CDH based ring signature which is one of the unforgeable and anonymous technique. No centralized entity is involved i.e., no concept of group manager. This scheme comprises two algorithms: Ring_sign and Ring_verify.
Ring-sign: This algorithm takes as input given message M. For a size of ring n, each group member chooses a secret key $Sk = x_i$ which belongs to $Z_p$ and public key $Pk = g^{x_i}$.

- Signer t uses global parameter $d$, $A_0$, $A_1$... $A_l$ $\in G_1$ of l random elements.
- Signer t will choose random $r_i \in Z_p$ for all other members of the group and generates $s_i = g^{r_i}$. Signer t again computes signature on behalf of group $s_t = (d \cdot \prod P_k^r_i (u_0 \cdot \prod A_j^{m_j})^{-r_{n+1}})^{1/\times t}$ (4)

In (4), signer t computes signature $s_t$ using his private key $X_t$ with different parameters such as: public keys of n group members PK, the global parameter $d$, $u_0, u_1$... $u_l$, message M divided into l elements ($M_1$...$M_l$), random number r. Final signature is $\sigma = (s_1, s_2$...$s_{n+1})$.

Ring-verify: As per (5), verifier verifies the signature using received ring signature $\sigma$, message M, and public keys PK of all members. The verifier checks the following equality.

$\prod_{i=1}^{n} e(s_i, P_k^r) (u_0 \cdot \prod A_j^{m_j})^{-r_{n+1}} = e(g, d)$ (5)

In (5), using public keys PK, signature $(s_1, s_2$...$s_{n+1})$, and received message M, recomputes $\prod_{i=1}^{n} e(s_i, P_k^r) (u_0 \cdot \prod A_j^{m_j})$. This recomputed part is verified using global parameter $d$ and $g$.

IV. PROPOSED SCHEME

A. System Model

The framework for our P2DPDP scheme is as shown in Fig. 2. It consists of three entities: Cloud User, CS, and TPA.

- **Cloud User**: Cloud user is one of the members of user group who can share a file in a group. Users can check the integrity of shared files through an audit log generated by TPA.
- **CS**: an entity having the capability of computation and storage at its end. It is having the responsibility to maintain and manage outsourced files.
- **TPA**: an external entity that works on behalf of users and expertise in verifying the integrity of outsourced data.

The workflow for P2DPDP scheme from Fig. 2 is as follows:

1) Any cloud user from a group outsources data file on CS. Before outsourcing, user calculates the tag for each block, signs the blocks using a ring signature scheme. 2) The user constructs RBMT tree using hash values of file blocks. User generates circuit for auditing program, obfuscates it, and sends to CS. Shares MAC key to TPA as well as sign a contract with auditor regarding verification activity. 3) Based on the frequency mentioned in the contract, TPA generates challenges and performs auditing activity. During an audit, using CDH based ring signature, TPA verifies group signature using public keys of all members in a group. 4) Generates log file for each activity. 5) Users can check the log entry at any time to verify the integrity of an outsourced file. 6) For dynamic updates, user sends the update command uc to the TPA. 7) TPA updates the RBMT tree according to uc and sends updated proof to CS for verification. If verification successful, CS sends signed proof to user. 8) User verifies proof and if successful, send updated data blocks ui to CS. CS updates data blocks accordingly.

B. Design Goals

To achieve privacy-preserving during integrity verification of outsourced data, proposed scheme P2DPDP should satisfy the following design objectives:

1) **Public verification**: to allow an external auditor to verify the integrity of outsourced data without downloading the data file.
2) **Privacy-Preserving**: data or user identity must not be revealed to TPA during auditing.
3) **Data Dynamic Support**: integrity verification process must support dynamic updations on outsourced data such as insert, delete and modify operations.
4) **Lightweight**: verification process must create minimum communication and computation overhead on user and TPA.
C. P$^2$DPDP Scheme

Our proposed P$^2$DPDP scheme works in four phases: Setup, Store, AuditData, and AuditLog.

**Setup:** Let G and $G_T$ be two multiplicative groups produced by g with order p contains bilinear map $e: G \times G \rightarrow G_T$. User $U$ selects a signing key pair ($ssk$, $spk$), $\alpha$, $v$ where $\alpha \rightarrow Z_p$ and $v = g^x \in G$. $U$ picks random elements $u_1, u_2$ .... $u_s$ and fixes pseudorandom permutation, function key $\pi_{key}(\cdot)$ and $f_{key}(\cdot)$ respectively. The secret and public parameters are $sk=\alpha, ssk$ and $pk=(v, spk, u_1, u_2$ .... $u_s)$ . Group members randomly select private key as $x_i \in Z_p$. Using key generation of CDH based ring signature scheme and $Pk_i = g^{x_i} \in G$ as a public key.

**Store:** Initially according to (6), $U$ divides the file $F$ into blocks $n$ and sector $s$ as in:

$$F = \{ f_i \} \forall i \leq n, 1 \leq s$$

(6)

**Tag Generation:**

$U$ chooses random element name for file and computes file tag as

$$\tau = name | n | u_1, u_2 \ldots \ u_s, x_i | gssk$$

and data tag as in (7).

$$\sigma_i = (H(i||name) \cdot \prod_{i=1}^{s} u_j^{\mu_i})^y, i \in [1,n]$$

(7)

In (7), $U$ calculates data tag for each block $i$ of file $F$ using random elements of each sector $u_1, u_2$ .... $u_s$ and hash value of block number and file name. Here $H$ is any secure hash function. $U$ generates processed data $M$ as $M = \{ M, \phi \}$ where

$$\phi = \{ \sigma_i, \tau \}_{i \in [1,n]}$$

**Constructing RBMT:**

$U$ first calculates hash values for each block of file $F$ using H2.

$$h_i = H2(m_i) \text{ where } 1 \leq i \leq n$$

Then generate tree $TR$ using RBMT on ordered hash values. Each leaf node $w_i$ stores the corresponding hash value $h_i$.

**Ring Signature Generation:**

$U$ has to sign a block on behalf of a group using CDH based ring signature scheme. $U$ randomly chooses $u_0, r_i \in Z_p$ and compute signature for all other group members except $U$ denoted by $j$ in (8).

$$S_i = g^{r_i} \text{ for } i \in \{ 2, \ldots , n+1 \}/j \}$$

(8)

Where, $n$ - number of members in a group

$j$ - serial number of the member in the signature who is signing it

$U$ computes signature for every member using respective random number $r$ of that user. Then computes $h=H(\phi||T)$ where $T$ is a timestamp. $U$ again computes $S_j$ using (4).

**Auditing Phase:**

$U$ outsources $M$ and $\phi^T$ calculated as in (9) on CS. In (9), $S_1, S_2, \ldots S_{n+1}$ is the signature generated for $n$ users by $U$ on behalf of group at time $T$.

**Outsourcing Auditing Task:**

During this task, $U$ chooses a MAC key $k$ and passes it to TPA using a secure network. $U$ also produces a circuit $Audit_k$ as described below.

$$Audit_k$$

Input: $\tau$, $\{ (i, v_i) \}, \mu_j \in [1, s] \cap \{ \cdot \}, v, g, spk$

Constant: MAC Key $K$

Verify $\tau$

If Invalid

Output 1

Else

Deconstruct $\tau$ to get $name, u_1, u_2 \ldots \ u_s$

If Ver $\{ ((i, v_i), \mu_j, u_j)_{i|\mu_j \in [1, s]} \sigma, name \} = 1$

Output MAC$_s (name || \{(i, v_i)_{i|\mu_j \in [1, s]}})$

Else

Output 1

This circuit is similar to auditing program which generate the MAC using embedded MAC Key $K$ based on given input. Uniform PPT algorithm iO proceeds with audit circuit $Audit_k$ as input and generates public parameter P as $P=\text{O}(Audit_k)$.

$$\text{Ver} \{ ((i, v_i), \mu_j, u_j)_{i|\mu_j \in [1, s]} \sigma, name \} \text{ in circuit denotes checking of (10)}.$$}

$$e(\sigma, g) = e(\prod_{i \in E} H(i||name)^{v_i} \cdot \prod_{j=1}^{s} u_j^{\mu_i}, v)$$

(10)

- $U$ outsources RBMT tree $TR$ with signature $S_igssk(TR)$ to TPA. If verification is successful, TPA accepts $TR$ else rejects considering the malicious author.

**Agreeing Parameters:**

- All group members need to sign on a public parameter $P_{pub}=(q, h_{root})$ where $q$ is the number of data blocks and $h_{root}$ is a Merkle root of $TR$.

**Contract $CT$ is established between $U$ and TPA as**

$$CT = \{ BI, Fr, b \}$$

$BI$ - block index from which auditing work will start.

$Fr$ - Frequency at which TPA launches a challenge.

$b$ - number of challenged data blocks for checking.

**AuditData Protocol:** This protocol mainly deals with checking the integrity of outsourced data and log generation by TPA.

**Auditing Phase:**

- TPA generates a challenge message $Q(b) = \{ b, K_1^{(b)}, K_2^{(b)} \}$ using data blocks $b$ to be audited. TPA
Generates pseudorandom permutation and function keys \( \{K_1(b), K_2(b)\} \) respectively and send to CS.

- CS computes \( i=\pi K_1(b)(\xi) \) and \( v_i= fK_2(b)(\xi) \) where \( \xi \in [1, b] \) and \( b \) is the size of I (input blocks to be audited).
  - Based on public parameters and corresponding \( \sigma_i \) calculated using (7), \( \tau, f_i, \) and \( b \), CS computes:

\[
\sigma(b) = \prod_{i \in I} \sigma_i^{v_i} \cdot \mu_i = \sum_{i \in I} v_i \mu_i
\]

- **Log Generation:**
  - \( PRF_{\text{new}} = PRF(b) \)
  - \( PRF_{\text{old}} = PRF(b) \)
  - **AuditLog Protocol:**
    - \( PRF_{\text{new}} \) is calculated by executing audit circuit \( \text{AuditLog} \) i.e. \( P=\text{Io}(\text{AuditLog}) \) using input parameters file tag \( \{i, v_i\} \) and \( \{v, g, spk\} \). CS sends this \( PRF_{\text{old}} \) and \( \text{Sig}_{gsk_{CS}}(PRF_{\text{old}}) \) to TPA for verification.
    - Using CDH based ring signature process, TPA verifies the group signature based on input signature \( \phi^T \), public keys \( \{PK_1, PK_2, \ldots, PK_n\} \) of all members in a group, \( F^T \) and public parameter \( u_0 \). TPA first calculate \( h=H(\phi||T) \). Then verifies signature using (5).
    - To verify the integrity of data, TPA computes \( i=\pi K_1(b)(\xi) \) and \( v_i= fK_2(b)(\xi) \) and compare MAC with \( PRF_{\text{old}} \) calculated by CS using (11).

\[
PRF_{\text{old}} = MAC_k(\text{name}||\{(i, v_i) \in I\})
\]

- **Log Generation:**
  - After every audit, either successful or fail, TPA creates a log record of his auditing work.

\[
L^{(b)} = \{t, Q^{(b)}, PRF_{\text{old}}, \text{Sig}_{gsk_{CS}}(PRF_{\text{old}})\}
\]

and saves in his local file \( \text{Log File} \).

- **AuditLog Protocol:**
  - TPA generates challenge using random subset \( B \) of file block indices and sends it to TPA. For each \( b \in B \), TPA finds challenge \( Q^{(b)} \), proof \( PRF_{\text{old}} \) from his log file. Computes \( i, v_i \) from \( Q^{(b)} \). TPA generates multi-audit proof \( \cup_p \) using \( h_{root} \) of TR and generates the proof of appointed log for subset \( B \) using (12).

\[
PRF_{\text{old}} = \{ \cup_p, i, v_i, PRF_{\text{old}} \}
\]

(12)

- In (12), \( B \) indicates the challenge generated by \( U \) during AuditLog. Elements \( i, v_i \) denotes the challenge retrieved through log file for blocks \( b \) and \( PRF_{\text{old}} \) indicates proof retrieved from log file for blocks \( b \). \( \cup_p \) is a multi-audit proof generated from RBMT.

  - TPA sends a signed proof with his signature \( \text{Sig}_{gsk_{TPA}}(PRF_{\text{old}}) \) to U for verification. After verifying the signature, U computes new PRF as.

\[
PRF_{\text{new}} = MAC_k(\text{name}||\{(i, v_i) \in CB\})
\]

- TPA compares if \( PRF_{\text{new}} = PRF_{\text{old}} \). If matched, verification of outsourced data is successful else verification fails. U also verify \( \cup_p \) using \( h_{root} \) of TR.

**D. Support for Dynamic Updates**

P\(^3\)DPDP scheme support three types of update operations such as: deletion, insertion, and modification on blocks. If we perform these updates one by one, it will incur a large computation overhead at the auditor side to generate and verify the hash tree. To reduce this overhead, P\(^3\)DPDP is based on a MLA scheme using RBMT proposed by ODPDP which can handle updates in batch instead of one by one.

Initially, \( U \) computes all the hash and tag values of the new file block in Store phase, generates the RBMT tree, and set public parameter as \( P_{pub} = \{q, h_{root}\} \). \( U \) sends the update command uc to CS and TPA. \( U \) also generates audit circuit same as basic scheme but with modified verification function VerD ((\( i, v_i, u_i, \sigma, \text{name} \)) \( \epsilon \) \( \{1, 8\} \)) which consists of checking following equation:

\[
e(\sigma, g)=e(\prod_{i \in I} H(\sum_{j=1}^{s} m_{ij} \cdot v_i), \prod_{j=1}^{s} u_j^{v_i}, v)
\]

After receiving uc command, TPA updates leaf nodes, other affected nodes and generate an updated tree TR’ and it’s Merkle root \( h_{root} \). TPA then sends the updated signed proof up to CS. CS verifies up by executing the audit circuit. \( U \) can also later check the correctness of up. If verification is successful, \( U \) sends updated information \( ui \) to CS and CS updates the processed data.

**V. Evaluation**

In this section, P\(^3\)DPDP scheme is evaluated by showing correctness proof, security analysis, performance and experiment analysis.

**A. Correctness Proof**

\( \text{AuditLog} \) is an audit circuit generated by \( U \) during Store phase. Upon execution of this audit circuit, CS generates MAC based on challenge, block tag \( \sigma \) and using global parameters. Audit circuit contains verification function which denotes the realization of (10). Correctness proof for (10) is as follows:

\[
e(\sigma, g) = e(\prod_{i \in I} \sigma_i^{v_i}, g)
\]

\[
e(\prod_{i \in I} (H(i||\text{name}) \cdot \prod_{j=1}^{s} u_j^{f(i) \cdot v_i}, g)
\]

\[
e(\prod_{i \in I} (H(i||\text{name}) \cdot \prod_{j=1}^{s} u_j^{f(i) \cdot v_i}, v)
\]

**B. Security Analysis**

- **Storage Correctness:**

  Theorem 1: If the CS successfully passes the verification from an auditor, then data outsourced on CS must be intact.

  **Proof:** Assume user \( U \) outsourced data at CS using Store protocol. But due to some problems, data at CS accidentally corrupted or deleted. With P\(^3\)DPDP scheme, malicious CS can't pass its verification. We prove this by game sequence as below:

  1. Based on a contract signed between \( U \) and TPA, TPA generates a challenge using AuditData Protocol.
  2. For the challenged blocks, CS computes \( i, v_i \). Using
3) Public parameters and \((k_1, k_2)\), CS computes \(\overline{PRF}\) and send to TPA.

4) TPA computes \(i = \pi \cdot k_1(\xi)\) and \(v_i = f(k_2(\xi))\) where \(\xi \in [1, c]\) and \(c\) is the size of \(I\) (input blocks to be audited).

5) Using \(MAC_k\), TPA calculates \(PRF\) and compares it with \(\overline{PRF}\).

6) CS wins if TPA passes the verification even if \(PRF \neq \overline{PRF}\).

But in above game, it’s very difficult for malicious CS to cheat auditor because of HMAC scheme during verification.

Liability:

Theorem 2: An honest auditor can demonstrate that he did his work correctly in case of any disputes.

Proof: To prove the liability of the auditor, we consider two situations: when an auditor is honest or an auditor is dishonest. Consider first the auditor is honest. As per the contract between auditor and User, the auditor generates a challenge \(Q^{(b)}\). User can reconstruct the challenge since the contract consists of number of data blocks to be audited. User can check the value of \(PRF\) by recalculating (11). Honest auditor generates a log file named Log_File which is the evidence for all the auditing work completed by auditor. So honest auditor can prove his liability by this Log_File.

Compare to this, if the auditor is malicious and not doing his work properly, user can use AuditLog Protocol to verify the behavior of auditor. By generating challenge, user can check the AuditLog file anytime and auditor can’t deny his misbehavior.

Privacy-Preserving:

Theorem 3: From the server’s response to the challenge message, TPA not able to infer any information such as data and identity of user.

Proof: During verification, user U first generates an audit circuit (which is nothing but an auditing algorithm program which is supposed to be originally executed by TPA). U obfuscates this circuit by embedding MAC key K and send to CS. For each verification, CS computes the inputs based on the challenge message and executes the obfuscated program. CS generates the MAC tag and sends it to TPA. TPA has to only verify the MAC tag to check the integrity of outsourced data. TPA needs to calculate \(i\) and \(v_i\) based on challenged blocks using the HMAC scheme. So, it is computationally infeasible for TPA to infer any information or user data using P\(^2\)DPDP.

P\(^2\)DPDP uses CDH based ring signature scheme to share any data among group members. In this scheme, user who want to share a file, computes signature on this data with his own private key using (4). During verification, TPA can verify the signature with public keys of all users. Using this scheme, TPA can check whether the signature is computed by a valid user of group or not but scheme can’t reveal individual user identity to verifier. Thus because of CDH based ring signature scheme and IO, P\(^2\)DPDP proved to be privacy-preserving.

C. Performance Analysis

We first evaluate the performance of P\(^2\)DPDP scheme which shows the privacy-preserving, lightweight auditing process. Also, we compare the performance of P\(^2\)DPDP with existing schemes.

The main important functionalities which we have considered for this work are: public auditing, dynamic data operations, privacy-preserving, and group support. Table I shows the functionality comparison of P\(^2\)DPDP scheme with existing schemes.

To evaluate the performance of P\(^2\)DPDP scheme, we evaluate the communication cost between CS and TPA during the proof generation and verification phase of AuditData protocol. Communication cost between the user and CS is not important since user uploads the data entirely to CS initially and user can verify the integrity of outsourced data during AuditLog protocol. During proof generation, TPA generates a challenge message \([k_1, k_2]\) for \(b\) number of blocks where \(k_1\) and \(k_2\) are transformed keys of HMAC. After receiving challenge message, CS generates proof \(PRF\) by calculating HMAC through the obfuscated program. So, communication overhead for proof generation is \(bH\) where \(H\) is any secure hash operation. After generating the proof \(PRF\) using HMAC, CS sends it to TPA. During verification, TPA checks the integrity of outsourced blocks using MAC key and verify

\[PRF = MAC_k(name||\{(i, v_i) | i \leq b\})\]

So TPA has to calculate only HMAC. TPA also verifies the user signature by verification algorithm of CDH based ring signature. Using (5), TPA verifies the users \(n\). So, the total communication overhead during verification is \(H_{2p} + n\) where \(H_{2p}\) is hashing operation into \(Z_p\).

User can check the integrity of outsourced data or performance of TPA during AuditLog protocol. So total Communication overhead during AuditLog is also \(H_{2p}\). Table II shows the comparison of P\(^2\)DPDP scheme with the existing scheme. Fig. 3 shows the comparison between ORUTA, CORPA, and our P\(^2\)DPDP scheme for communication overhead in KB with respect to group size.

We can’t compare this cost with ODPDP scheme since it doesn’t support user groups. The comparison demonstrates the noticeable and constant performance of communication cost between CS and TPA during auditing in P\(^2\)DPDP scheme.
TABLE I. THE COMPARISON OF AUDITING FUNCTIONALITIES

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Third-Party Auditor</th>
<th>Dynamic Data Operation</th>
<th>User Group Support</th>
<th>Privacy Preserving</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORUTA [5]</td>
<td>Homomorphic Authenticators (HA)</td>
<td>Index Hash Table</td>
<td>Ring Signature</td>
<td>Homomorphic Authenticable Ring Signature (HARS)</td>
</tr>
<tr>
<td>Tian et al. [20]</td>
<td>Homomorphic Verifiable Authenticators (HVA)</td>
<td>Dynamic Hash Table</td>
<td>Group Signature</td>
<td>Random Masking</td>
</tr>
<tr>
<td>ODPDP [21]</td>
<td>Efficient HVA</td>
<td>Rank Based Merkle Tree (RBMT)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>CORPA [23]</td>
<td>HA</td>
<td>Merkle Hash Tree (MHT)</td>
<td>Group Signature (GS)</td>
<td>HA+GS</td>
</tr>
<tr>
<td>Thokcham and Saikia[34]</td>
<td>Vector Commitment</td>
<td>MHT</td>
<td>CDH based Ring Signature</td>
<td>CDH based Ring Signature</td>
</tr>
<tr>
<td>Zhang et al. [36]</td>
<td>Indistinguishability Obfuscation (IO)</td>
<td>MHT</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Proposed Scheme P^2DPDP</td>
<td>Indistinguishability Obfuscation (IO)</td>
<td>Rank Based Merkle Tree (RBMT)</td>
<td>CDH based Ring Signature</td>
<td>CDH based Ring Signature</td>
</tr>
</tbody>
</table>

TABLE II. COMMUNICATION COST

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Proof Generation</th>
<th>Proof Verification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AuditData</td>
</tr>
<tr>
<td>ORUTA [5]</td>
<td>((s+nb)E+nbM+bsM+sH)</td>
<td>((2s+b)E+(2s+b)M+nM+bH+(n+2)p)</td>
</tr>
<tr>
<td>ODPDP [21]</td>
<td>(2q+s+1)E</td>
<td>((b+s+1)E+2p)</td>
</tr>
<tr>
<td>CORPA [23]</td>
<td>(H+M)</td>
<td>(2E+P+bE+bM)</td>
</tr>
<tr>
<td>Proposed Scheme P^2DPDP</td>
<td>(bH)</td>
<td>(H+n)</td>
</tr>
</tbody>
</table>

\(s\): Total no. of sectors  
\(n\): Total no. of users in a group  
\(b\): No. of challenged blocks  
\(q\): Total no. of blocks  
\(E\): Exponentiation operation  
\(M\): Multiplication Operation  
\(H\): Hash Operation  
\(P\): Pairing Operation

Fig. 3. Communication Cost w.r.t. Group Size.

D. Experimental Results

This section proves the performance of P^2DPDP system in terms of different experiments. We deployed our P^2DPDP scheme on a system comprising Windows 8.1 with an Intel Core i5-5200U CPU functioning at 2.20 GHz, 4.0 GB RAM. Python is used for module implementation of P^2DPDP scheme. The hash algorithm is instantiated using SHA256. Fig. 4 shows the impact of number of users in a group on verification time during AuditData protocol. It shows that verification time is independent of the number of users. By creating a group of 25 users, we have compared the results with Oruta and CORPA scheme. We have not considered ODPDP scheme for comparison since it doesn’t support user groups. All results are average of 5 runs. For 20 users in P^2DPDP scheme, the Verification time is 0.15 seconds. While Oruta, and CORPA are 2.24, and 1.75 seconds respectively. Result proves the effectiveness and lightness of our scheme because of reduced and constant verification time at auditor side during AuditData protocol. Since our P^2DPDP scheme is based on ODPDP scheme, it is mandatory for us to compare results with this scheme. In both the schemes, verification is performed during AuditData and AuditLog protocol. Initially we compare the results of both schemes during AuditData where TPA performs audit based on contract and generate log entries. Fig. 5 shows the proof generation and verification time in seconds with respect to challenged data blocks during AuditData protocol.

To compare the results with ODPDP scheme, in P^2DPDP scheme, the block size is kept fixed i.e., 16KB. Total outsourced data is 1GB. Fig. 5(a) shows the constant proof generation time in P^2DPDP scheme as compared to ODPDP. Fig. 5(b) shows the gradual increase in proof verification time as number of challenged blocks is increasing in P^2DPDP as compared to ODPDP scheme. Fig. 6 shows the performance of P^2DPDP during AuditLog. It presents the computation time and communication cost required by user to verify the past work of TPA under the number of checked log entries.
The paper proposes the P₂DPDP scheme for cloud storage in which TPA are actively involved during verification process which generates challenges based on the contract signed between TPA and user. TPA also generates log which can be audited by user as per his convenience. P₂DPDP scheme also create the lightweight verification process so as to reduce the computation burden of TPA using new cryptographic technique, Indistinguishability Obfuscation and MAC. P₂DPDP support and manages user groups using CDH based ring signature scheme. CDH based ring signature is anonymous scheme which preserves the identity of users from TPA during auditing. P₂DPDP scheme supports dynamic updates in batch mode using MLA solution proposed by ODPDP scheme which is based on RBMT.

Security analysis and experiments show that P₂DPDP scheme is secure, lightweight and privacy-preserving. Communication cost during auditing between CS and TPA is almost constant and reduced compared to Oruta and CORPA since TPA has to just calculate MAC and compare it with MAC received from CS. Verification time is also reduced and constant compare to existing schemes. Experimental results reveal that verification time is independent of number of group users. Results of AudiLog protocol shows that P₂DPDP scheme is performing better in terms of communication time and cost as compared to ODPDP scheme. CDH based ring signature generates certificates which need to be processed during verification leads to increase computation time. In terms of future work, we plan to modify P2DPDP scheme using certificateless signature schemes to reduce computation time.

VI. CONCLUSION

In most of the previous auditing scheme, Cloud user and TPA are actively involved during verification process which may create additional burden in terms of time and cost. This paper proposes P₂DPDP scheme for cloud storage in which there is no need of user during verification process. TPA generates challenges based on the contract signed between TPA and user. TPA also generates log which can be audited by user as per his convenience. P₂DPDP scheme also create the lightweight verification process so as to reduce the computation burden of TPA using new cryptographic technique, Indistinguishability Obfuscation and MAC. P₂DPDP support and manages user groups using CDH based ring signature scheme. CDH based ring signature is anonymous scheme which preserves the identity of users from TPA during auditing. P₂DPDP scheme supports dynamic updates in batch mode using MLA solution proposed by ODPDP scheme which is based on RBMT.

Security analysis and experiments show that P₂DPDP scheme is secure, lightweight and privacy-preserving. Communication cost during auditing between CS and TPA is almost constant and reduced compared to Oruta and CORPA since TPA has to just calculate MAC and compare it with MAC received from CS. Verification time is also reduced and constant compare to existing schemes. Experimental results reveal that verification time is independent of number of group users. Results of AudiLog protocol shows that P₂DPDP scheme is performing better in terms of communication time and cost as compared to ODPDP scheme. CDH based ring signature generates certificates which need to be processed during verification leads to increase computation time. In terms of future work, we plan to modify P2DPDP scheme using certificateless signature schemes to reduce computation time.

REFERENCES


