Information Systems and e-Business Management (2020) 18:351–377
https://doi.org/10.1007/s10257-019-00417-8

ORIGINAL ARTICLE

Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework

Indumathi Jayaraman[1] · Mokhtar Mohammed[1]

Received: 16 May 2019 / Revised: 24 May 2019 / Accepted: 5 July 2019 / Published online: 2 August 2019

Abstract

The assiduous parade of the state-of-the-art sprouting digital technologies, is disrupting the smooth, easy-going health care digital ecosystem and forewarns us to manage it preemptively; since adaptation and survival of the fittest is a proven fact and we need to acclimatize to the mutated health care digital landscape. In this paper, the heightened consternations in the cloud are discoursed, with prime focus on integrity and privacy solutions, useful to hook the doles of cloud computing technologies for the health care world. An all-embracing appraisal of the correlated up-to-date research work on Provable Data Possession (PDP), tosses light on the erstwhile current status, research challenges, and future directions of PDP based health care data integrity. The need of the hour is a system, which, aids as an external auditor to audit the user's outsourced health care data in the cloud, deprived of the wisdom of the health care data content. The contributions in this paper are (1) A comprehensive analysis of the contemporary Privacy Conserving PDP data integrity schemes, (2) a proposed novel generic support framework, which is useful to shield stored health care data, provide authentication in the cloud environment, which, is scalable and efficient, (3) deployment of the Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework. The results validate that the proposed SPC-PDP framework can competently accomplish secure auditing and outclass the erstwhile ones. The SPC-PDP framework is no doubt, a promising solution to the challenges soaring due to the state-of-the-art improvements in health care digital technology. Last but not the least, this paper also gives a bird's eye view on the future directions of secure and privacy preserving data integrity.

Keywords Authentication · Availability · Cloud computing · Confidential · Data dynamics · Data privacy · Health care data · Integrity · Merkle Hash Tree · Non-repudiation · Privacy preserving · Provable data possession · Storage correctness · Secure

✉ Indumathi Jayaraman
indumathi@annauniv.edu; indumathi.j@gmail.com

Extended author information available on the last page of the article

> *"Everyone designs who devises courses of action aimed at changing existing situations into preferred ones".*
> *Herbert Simon.*

1 Introduction

The inescapable pageant of new technologies is unfolding on many fronts. "Cloud computing", like a Tidal wave has pranced every sphere of life.

Cloud technology demonstrates to be a great asset if chosen for any realm, as it is more *cost effective*, *quick information deployment, less management*, *environment friendly*, *globalization of work & streamline workflow*, *infinite storage capacity, lesser personal training cost, more flexible, more scalable*, *reduced usage of bandwidth*, and *usability*.

The potential downsides of cloud are loss of *control over the cloud, Outages* and data losses (e.g. EC2, Amazon's cloud-computing platform, underwent three major outages in the past 3 years, and lost some customers to lose their data permanently. Others include the Netflix, Pinterest, Airbnb and Instagram), *Privacy, Integrity and Security.* Among the downsides the cloud security issues, integrity tops the list.

Almost all the evolving technologies modify the professional or social landscape disrupting the status quo, altering the way people go for health care and treatment, and rearrange value pools. Health care digitization has drifted the Homosapiens from an information scarce ecosphere to an information rich biosphere. The health care data commodity spawns not only a lucrative, fast-growing health care digital industry, prompting security and privacy regulators to step in, to restrain the threats and vulnerabilities that hamper its utility.

The implications of that shift, has also pressurized the research community, to find novel ways for safeguarding the data currency fueling the health care digital information age. This research paper provides a comprehensive study for health care data security with special emphasis on the health care data integrity in cloud computing. This paper analyses and gives a cloud integrity (security) attention seeking syndrome, haunting the health care digital world, which has to be immediately plugged.

The key contributions of this paper are summarized as follows:

a. A comprehensive analysis of the contemporary Privacy Conserving PDP health care data integrity schemes.
b. Propose a novel generic support framework, which can concurrently support the vital functions, like, Privacy Conserving health care data storage, integrity, batch auditing and dynamic data auditing. The proposed framework is used to provide safe storage, provide authentication, and to scrutinize the health care data in the cloud environment, which, is scalable and efficient. This architecture also supports health care data authentication and confidentiality for cloud storage.
c. The Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework, deployment results validate that the SPC-PDP framework, can competently accomplish secure auditing and outclass the erstwhile ones, in terms of scalability,

public auditing, batch auditing, efficiency, reliability, privacy, cost-effective computational overhead and communication costs. The SPC-PDP architecture is no doubt, a promising solution to the challenges soaring due to the state-of-the-art improvements in digital technology.

The rest of the paper is organized as follows. In Sect. 2, the background, motivation for this paper, cloud computing issues, and cloud storage issues with special emphasis on cloud security issues are discussed. Section 3 discusses the overview of all related works, and existing systems. Section 4 formulates the problem statement, and description of the proposed problem and its solution. Section 5 deals with the proposed Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework and its detailed design. Section 6 explains the comprehensive implementation of modules, setup, and the pseudo code. Section 7 deals with the experimental evaluation of the SPC-PDP framework, observed result and analysis. Section 8 highlights the conclusions. Section 9 describes the future work.

2 Background

The contemporary epoch's most intriguing paradigm, which, is hastily progressing and emerging, at lightning speed, it becomes critical to understand all aspects about this technology.

2.1 Need for outsourcing health care data

The varied health care data are created at high velocity, and it is voluminous, making it unsurmountable to store in local machines. Subsequently, the need for storing and sharing of health care data, has grown leaps and bounds. Hence, it becomes forceful to have, all the records of the patients to be outsourced for future references in the Cloud. Outsourcing data to the cloud has become one of the most popular applications in cloud computing as SaaS paradigm (Buyya et al. 2009; Wang et al. 2016b). Outsourcing data to the cloud offers reimbursements like Cost Reduction (Sookhak et al. 2015), easy infrastructure management (Kuo 2011), availability and scalable healthcare services (Sookhak et al. 2015).

The EHRs need to be shared among the various care delivery organizations and multiple healthcare providers (Zhang and Liu 2010) for the purpose of consultation and treatment. The archetypal information in a PHR are treatments and diagnosis, surgeries, laboratory reports, insurance claims data, personal notes and wellness charts that the patients use in order to keep track of their health (Li 2013). The records are to be accessed globally to support patient needs, improve the quality of service, accelerate biomedical discoveries, reduce medical costs, and well-timed decision making. The health care data hoarded in the cloud is used to augment collaboration among innumerable partaking entities to the healthcare domain (Ahuja et al. 2012), and to offer the facilities like scalability, agility, cost effectiveness, and round the clock availability of health-related information (Abbas et al. 2015; Wu et al. 2012). The cloud stored patient information is exploited for the prosperousness of the community, medical diagnosis, and other health-related discoveries. But the health care data owners (DO) are underprivileged of their direct control over health care data, which makes them vulnerable to various security threats. Since, the sensitive electronic health data and personal health information (PHI) is stored and shared in the cloud, various privacy and security concerns arise (Li et al. 2010). The patient's medical history as well as the doctor's activities is hunted down to privacy and confidentiality challenges.

2.2 Cloud computing

The definition issued by the U.S. National Institute of Standards and Technology (NIST) September, 2011 defines, *"Cloud computing is a 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. This cloud model is composed of five essential characteristics, three service models, and four deployment models"*. (The NIST Definition of cloud computing, SP800-145.pdf).

The NIST definition has been accepted as the "defacto standard" (Bernd Grobauer et al. 2011), by researchers. The key security aspects that is anticipated of any health care data sharing technology are Confidentiality, Integrity, Authentication, Authorization, Non-repudiation and Availability. Despite the doles offered by cloud computing (based on 3-4-5 rule), and the fact that many enterprise applications and health care data are moving into cloud platforms; one should not forget that it still has security issues which is a major barrier for cloud adoption (Bansidhar et al. 2011).

2.3 Cloud security issues

Amongst the many issues that demand attention, from the research community and make them feel overwhelmed in finding solutions are *Elasticity, Insecure APIs, IP address, Network Insecurity, Provider Security Malfunction, Reliability, Availability, Virtualization, Confidentiality and Privacy, Integrity, Malfunction Time, Data Location, Native Customer Attacks.*

Amongst the five pillars (integrity, authentication, availability, confidentiality and nonrepudiation) of information assurance (IA), in sharing health care data/information, Integrity tops the list. The subject of guaranteeing health care data integrity over the remote servers has been debated for numerous years; several solutions have been suggested to tackle this problematic issue. This paper is aimed to concentrate on how to apply health care data integrity to massive health care data stored in the cloud.

2.4 Threats arising due to health data privacy

Threats to Health Data Privacy in the Cloud include the *Spoofing Identity* (unlawful attempts by other users, or machines, to pose as the valid users or machines (Metri and Sarote 2011), *Data Tampering* (malicious attempt to modify the data contents is called data tampering (Zissis and Lekkas 2012), *Repudiation* (denying the obligations of a contract (Chen et al. 2012), Denial of Service (DoS) [services are denied to privileged users (Lounis et al. 2015), *Unlawful Privilege Escalation* [Unlawful users may obtain access to data and can subsequently infiltrate into the system, such that the data contents at a large-scale are compromised (Metri and Sarote 2011).

2.5 Challenges arising due to features of cloud computing

One can list below few of the issues, that, are arising due to the features of cloud computing

- *Outsourcing*: As the data is outsourced for data storage to a third party, control over the data is lost, leading to security and privacy challenges in cloud environments. This can be solved by protecting, controlling and verifying to ensure their confidentiality, integrity, and privacy along with cloud reliability, availability and service continuity (Xiao and Xiao 2012; Tang et al. 2016).
- *Multi-tenancy*: The data stored in a single location may be owned by diverse users leading to new threats, arising out of co-residence issue and its related attacks.
- *Colossal data*: The exponentially exploding volume of data and applications, brings new challenges to upkeep dynamic data monitoring, privacy preservation and security protection. As the existing security mechanisms are unable to handle the dynamic data patterns, attributes and access rights, new strategies and protocols are to be developed (Xiao and Xiao 2012; Rong et al. 2013).

2.6 Security and privacy requirements

Amidst, the many issues that demand attention from the research community and urge them to find appropriate solutions (*and from the perception of a client*), three grave requirements are to be contented to pass the security test of their outsourced data contents:

- *Data confidentiality*: A fine-grained access control mechanism of granting privileges to the outsourced health data must be protected from the external entities, such as the CSP, and from the unauthorized insiders (Abbas and Khan 2014).
- *Data integrity*: The outsourced data is to be correctly and trustworthily stored without tampering, ensuring the accuracy and completeness of data.
- *Privacy conservation*: User demand is to protect personal data information, and safeguard the unlink ability (identity protection, personal contents) between dif-

ferent accesses to outsourced data. The *requirements for privacy conservation* to be fulfilled are collusion resistance, anonymity, authenticity, and unlinkability.

Having comprehensively discussed the numerous issues apprehensive for data security in the cloud (Bhadauria et al. 2011; Liu 2012); and spiraled by the human instinct of looking at the problems as wonderful opportunities to find the appropriate technical solutions, this paper is aimed at a trifling step. Security and privacy are like the two eyes and they are indispensable. Several mechanisms and the relevant concepts of preserving privacy are existing in literatures (Gitanjali et al. 2007, 2008, 2009a, b; Indumathi 2012, 2013a, b, c; Indumathi and Uma (2007a, b, 2008a, b, c, d; Murugesan et al. 2009, 2010a, b; Prakash et al. 2009; Satheesh Kumar et al. 2008; Vasudevan et al. 2007).

3 Literature survey

To ensure privacy and integrity of the data, various integrity checking techniques has been proposed from time to time. Let us take a look at the comprehensive survey of different integrity checking techniques with their advantages and disadvantages. A few of the "*privacy proving*" techniques devised as technology solutions in yester years are *encryption, de-identification, identity-based anonymization and suppression.*

Ateniese et al. (2009) proposed, *Data Protection As A Service,* which, provides both data security and privacy; and an evidence of privacy for data owners in the presence of potential threats. A *partially dynamic PDP* scheme, proposed by Ateniese et al. (2008) upkeeps the data dynamics. Sebé et al. (2008) proposed protocol that aids in unlimited times of file integrity verifications and allows predetermined compromise between the protocol running time and the local storage burden at the user.

Erway et al. (2015) constructed a fully data dynamic auditing scheme, using a *checklist.* Wang et al. (2010) proposed the *Privacy-Preserving Public Auditing* of stored data, in which, the TPA is used to efficiently and simultaneously perform data audits for multiple users. Due to the large number of data tags, their auditing protocols incurs a heavy storage overhead on the server The Wang et al. (2010) proposed a *Privacy-Preserving PDP*, wherein a *Homomorphic authenticator is integrated with random masking* to achieve privacy preserved integrity and supported public auditing. It did not support data dynamics. Wang et al. (2011) used *Merkle Hash Tree* to construct a remote data integrity auditing scheme that aids full data dynamics. Ren et al. (2012) proposed a *security protocol* (*Privacy As A Service*) to provide security and privacy feedback for the client when storing and retrieving data.

The Cooperative Provable Data Possession (CPDP) proposed by Zhu et al. (2012) used *Hash Index Hierarchy* to achieve privacy preserved integrity, to support public auditing and batch auditing in multicloud. It does not support dynamic audit and auditing for multiuser. Under the proposition, the TPA is a trustworthy agent, a data owner outsources data to the remote cloud or delegate the auditing task to the third party. Obviously, this is not a logical assumption increasing the possibility of leaks.

Data owner does not want the TPA, of course, to know the details of data contents (Mandagere et al. 2008; Meyer and Bolosky 2012). How to effectively solve the problem of the development of the technology of the RDA is very imperative? To protect the identity privacy of user, Wang et al. (2012) designed a privacy-preserving shared data integrity auditing scheme by modifying the ring signature for secure cloud storage.

In order to protect the data privacy, Wang et al. (2013) projected a *privacy-preserving remote data integrity auditing scheme, named as Oruta*, with the employment of a random masking technique. Ren et al. (2013) designed a Designated-Verifier Provable Data Possession (DV-PDP) protocol, in which, the data owner anticipates a particular user to authenticate the files in cloud storage. The DV-PDP is insecure against replay attack propelled by the malevolent cloud server. The DAP proposed by Yang and Jia (2013) used *index table* to achieve privacy preserved integrity, to support dynamic auditing and batch auditing in multicloud. It resulted in high computation cost.

Wei et al. (2014) proposed *SecCloud*, a privacy cheating deterrence and secure computation auditing protocol to avert an adversary from accomplishing a sensitive cloud data by using designated verifier signature schemes (Huang et al. 2011a, b; Zhang and Mao 2008). Commitment-Based Sampling (CBC) technique, used in the conventional grid computing (Du et al. 2004) diminishes the computational and communication overhead; SecCloud provides the batch verification for multi-users by using identity based aggregate signatures. Worku et al. (2014) used a *random masking technique* to improvise the efficiency of Oruta and to create a remote data integrity auditing scheme supporting data privacy protection. Guan et al. (2015) proposed remote data integrity auditing scheme (works on the *indistinguishability obfuscation technique*) minimizes the computation burden of signature generation on the user side.

Wang et al. (2015) used a *proxy re-signature* and proposed a shared data integrity auditing scheme with user revocation. Luo et al. (2015) used the *Shamir secret sharing technique*, and constructed a shared data integrity auditing scheme supporting user revocation. Subsequently, in order to overcome the considerable overheads arising due to the complicated certificate management (due to the use of Public Key Infrastructure (PKI)), many techniques were proposed. One such solution is the *identity-based remote data integrity auditing scheme in multicloud storage* proposed by Wang et al. (2015) *simplifies the certificate management* and used the user's identity information such as user's name or e-mail address to replace the public key.

Tian et al. (2017) proposed a *DHT-PA (Dynamic hash table-public audit) that uses a dynamic hash table*, to support public auditing, Privacy preserving, Support dynamic auditing and Batch auditing in multi cloud. The Communication cost is greater than DAP and IHT-PA. *A proxy is introduced to process data for users* in the novel identity-based proxy oriented remote data integrity auditing scheme proposed by Wang et al. (2016a). Yu et al. (2017) constructed a remote data integrity auditing scheme with perfect data privacy preserving in identity-based cryptosystems.

Wang et al. (2016b) proposed an *identity-based data integrity auditing scheme* satisfying unconditional anonymity and incentive. Hitachi has recently advanced technology to anonymize Encrypted personal data (2016), "Enigma: decentralized computation platform with guaranteed privacy", by Zyskind et al. (2015). To lessen the loss of users' key exposure, Yu et al. (2015, 2016) and Yu and Wang (2017) projected key-exposure resilient remote data integrity auditing schemes based on key update technique (2018). Yang et al. (2016) created an efficient shared data integrity auditing scheme, to support the *identity privacy and identity traceability of users.* Fu et al. (2017) used a *homomorphic verifiable group signature* to design the privacy-aware shared data integrity auditing scheme. Shen et al. (2017) introduced a *Third Party Medium (TPM),* which helps user generate signatures in the light-weight remote data integrity auditing scheme to provide data. Zhang et al. (2018) designed an *identity-based remote data integrity auditing scheme* to aid the real efficient user revocation. To overcome the limitation of handling only one update, per data block at a time He et al. (2018) proposed a method that can have simultaneous updates of multiple data blocks, by using the balanced data structure SGMHT an extension of the Merkle Hash Tree (MHT). This is founded on scapegoat tree, and it employs an erasure-coded hierarchical log structure to back the delayed update of multiple blocks and the data retrievability. The RDPC scheme proposed by Yan et al. (2019), overcomes the problems of Ren et al. (2013), by making the data owner to specify a unique verifier to check the data integrity. The RDPC protocol is grounded on the computational Diffie–Hellman assumption, and security is proved by the RDPC scheme in a random oracle model. The theoretical analysis and experiment results of RDPC scheme has less communication, storage, and computation overhead while achieving high error detection probability. The deterministic private verification data integrity check scheme proposed by Khedr et al. (2019) efficiently provides integrity and possession guarantees and works by utilizing the modified RSA-based cryptographic accumulator to confirm the integrity of the outsourced data.

To state a few of other notable research works embarked in this field (in chronological order) are the, "Enhancing cloud security using Data anonymization", by Sedayao (2012). "A Precautionary Approach to Big Data Privacy", by Narayanan et al. (2016), "Big data privacy: a technological perspective and review", by Jain et al. (2016).

Let us also look at a few of other notable research works to be analyzed. Remote integrity involves verification of data using a third-party. Provable Data Possession (Gasti et al. 2010; Wang et al. 2011; Juels and Kaliski Jr 2007) technique involves a client machine to verify remote data without downloading it. This technique uses the probabilistic possession of a random dataset from the remote server with the aid of homomorphic linear authenticators. The client must access the complete data block, to realize deterministic verification. Wang et al. (2011) used tree data types for block-tag authentication and protracted the work on proof of storage for data dynamics and attained public auditability for dynamic data operations and blockless verification.

To overcome the challenge of facing difficulties in verifying small data updates; techniques like ranked Merkle Hash Tree Wang et al. (2009), extended for the cloud-auditing scheme. Ateniese et al. (2009), technique maintains data integrity based on the signature scheme and provides authorized auditing. However, all of existing remote data integrity auditing schemes cannot support data sharing with sensitive information hiding. In this paper, we explore how to achieve data sharing with sensitive information hiding in identity-based integrity auditing for secure cloud storage.

3.1 Literature survey on existing approaches to retain integrity of health care data in the cloud

The different approaches that have been used from time to time to ensure the integrity of health data in the cloud computing environment is presented here in this section.

Mashima and Ahamad (2012) approach enhances the integrity and accountability of the EHRs by enforcing either the explicit or implicit patient control over the EHRs, by using the PKE to encrypt the health records. However, the assumption results in information disclosure as a result of any malicious activity by the issuer's. Wang et al. (2014a, b) approach to safeguard the health data is to deploy an independent third-party to maintain health data integrity and SKE encrypted data is uploaded. Homomorphic encryption (Lin et al. 2013) was used with IBE to preserve patients' privacy in the context of a mobile health monitoring system.

The Yang et al. (2015) model uses cryptography and statistical analysis to offer multi-level privacy. A limitation is that the data recipient, can act malevolently and reveal the information that can help linking the portions of patients' medical records. Lounis et al. (2015) has proposed a secure and scalable cloud-based architecture for medical wireless networks and it offers integrity of the outsourced medical data and a fine-grained access control is implemented through the CP-ABE based construction. This approach has to come out of the management issues rising due to the common policy changes, predominantly in the case of access revocation.

The other strategies used to maintain the integrity of health data in a cloud environment are PKE and digital signatures (Kaletsch and Sunyaev 2011); Hierarchical Predicate Encryption (HPE), ABE (Akinyele et al. 2010) and policy-based authorization methodology (Haas et al. 2011).

3.2 Analysis of some existing systems and expected solutions

Nevertheless, all of the prevailing remote data integrity auditing schemes in the literature, it is found that the existing systems have the following research gaps like *Cloud users are not able to safeguard and know the status of their outsourced data status, Extravagance in Input/output and transmission cost, quick to limp back*, *Late to recover the data loss or damage, Unapproved leakage of data and Privacy.*

The pressing need of the hour is a public auditing system of data storage security and a privacy-preserving auditing framework, which can tackle the evolving digital cloud storage technologies and aid an external auditor to audit user's outsourced data in the cloud deprived of the wisdom on the data content.

4 Problem statement

Propose, design a generic support framework that will safeguard storage, authenticate, audit, validate data in the cloud environment; which is also scalable, efficient as a public auditing system, and can accomplish batch auditing. The framework christened as, "Generic Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework", has to concurrently support data authentication, verification, auditing, integrity, and confidentiality for cloud storage.

5 Proposed Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework

The requirements that must be achieved, to create a proficient PDP framework, has been discussed by several authors (Yang and Jia 2012; Worku et al. 2012; Wang et al. 2013) in the preceding years. Upon a thorough study of the research papers the prerequisites that are desired for a good framework is comprehended as *Storage correctness/Unforgeability* (ensures that the cloud cannot cheat and pass the auditing process without storing the data properly), *Low storage cost, Low communication cost between the parties, Low computation cost* (low complexity of the computation), *Unbounded number of audits*, *Recoverability*, *Public auditing* (enables the TPA to check the integrity of the stored data in the server on behalf of the user), *Batch auditing* (lets the TPA to accomplish multiple auditing tasks at once from different Users), *Blockless verification* (the auditor must not retrieve any data blocks for the duration of the auditing process), *Stateless verification* (the TPA is to maintain and update state between the auditing process), *Privacy-preserving* (safeguards the privacy of the stored data content), *Dynamic data* (aid in dynamic data operations with less operating cost).

Figure 1 shows the High Level Design of the (framework of the) proposed Secure Privacy Conserving Provable Data Possession (SPC-PDP), to store the health care data in the cloud in a much secured manner; without the TPA unaware of the contents of health care data.

Provable Data Possession (PDP) mechanism efficiently audits and verifies the integrity of data held by unreliable third parties, like cloud storage service providers. Although multiple PDP schemes have been proposed/developed, no PDP schemes have been implemented with an existing cloud service and there is no research to date that delivers in-depth analysis on scalability, public auditing, batch auditing, efficiency, reliability, privacy, cost analysis for PDP. This research fills that gap by collecting and analyzing cost data for four PDP schemes, providing generic cost models (mathematical formulae expressing abstract models which can be used to infer future cost), and comparing the overall cost efficiency of each PDP scheme. The proposed framework is the need of the hour, in order to overcome the existing limitations of the PDP.

The SPC-PDP framework uses the generic Provable Data Possession (PDP) mechanism. It has four entities, as show in Fig. 1.

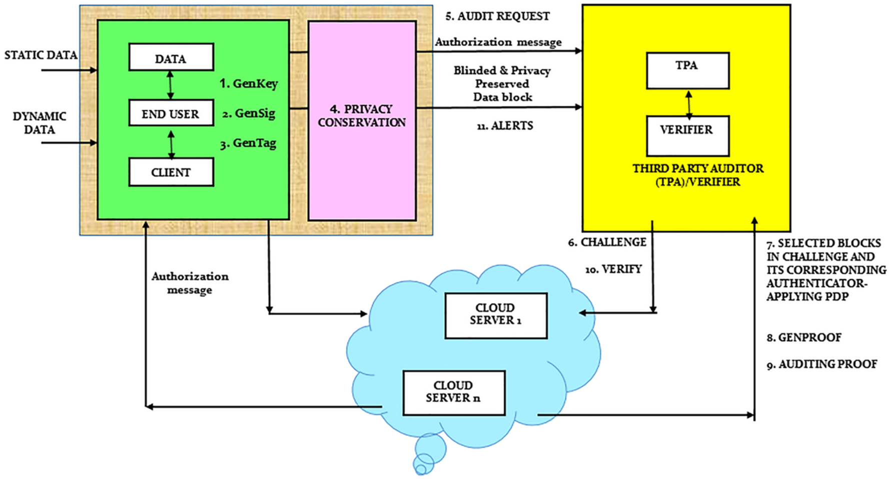


Fig. 1 Proposed Secure Privacy Conserving-Provable Data Possession (SPC-PDP) framework

1. *End user (CU)[client/cloud user/data owner/customer/consumer* who has data files to be stored in cloud], who comprehends the span of data files to be stored in the cloud.
2. *Cloud server (CS)* managed by the Cloud Service Provider (CSP) for providing data storage service and has significant storage space as well as computation resources.
3. *Privacy conservation* preserves privacy of the data. There are various Privacy Conserving techniques like aggregation, sampling, perturbation, and sanitization. In this project, the sanitization techniques are used to sanitize the health care data blocks matching to the sensitive information in the file. Summing up, the health care data blocks signatures are transformed into valid ones for the sanitized file, and the sanitized file and its matching signatures are uploaded to the cloud. The proposed SPC-PDP framework is designed and tested to handle any one of the following Health care data Sanitization techniques like, Nulling Out, Masking data, Substitution, Shuffling Records, Number Variance, Gibberish Generation and Encryption/Decryption.
4. *Third party auditor (TPA)/verifier* who has expertise and capabilities that cloud users does not have and it is trustful for evaluating the cloud storage service reliability on behalf of the user request.

6 Implementation

In this section, we endorse a comprehensive implementation, testing and experimental evaluation of the SPC-PDP framework on real-world data sets. In this research work the generic Provable Data Possession PDP is used on this framework. Based on the scope proposed in the architecture, the various techniques that have been deployed over the yester years are also tested.

Table 1 Notations

Symbol	Description	Symbol	Description
b	File block	$\boldsymbol{\Phi}$	Corresponding signature set
b_{size}	Block size in bytes	p_k	Public key
C	Client	S	Server
$c_0; c_1;: : :,$	Model-specific constants	s_k	Private key (secret key)
chal	Challenge	$\mathbf{s_{kID}}$	User's private key
E	Ordered collection of the corresponding verification metadata Tb	s_{size}	Sector size in bytes
F	File/ordered collection of all file blocks b	$\mathbf{s_{sk}}$	User's signing private key
$\mathbf{F}^*$	Blinded file	Tb	Verification metadata
f_{size}	File size in bytes	V	Proof of possession V for the blocks in F that are determined by the challenge chal
k	Security parameter	Σ	Ordered collection of the corresponding verification metadata Tb
name	File identifier name	τ	File tag

The elucidation for competent systematization and sustain of facets, analysis, and diverse user groups are based on a common conceptual Provable Data Possession. One of the technique tested on this framework is the *public key-based homomorphic linear authenticator (HLA)* and random masking. The *Privacy-Preserving Public Auditing* (PP-PDP) technique, proposed by Wang et al. (2010, 2013) uses the public key-based homomorphic linear authenticator (HLA) integrated with random masking. It drastically diminishes the communication and computation overhead as TPA does not need to have a local copy, and assures the data privacy without TPA learning the contents. The batch auditing is provided by the aggregation and algebraic properties of the authenticator. In short, HLA permits any person to certify the output of a complex computation over a large authenticated data with only a short tag.

PP-PDP I & PP-PDP II

- Data auditing is performed using HLA with random masking.
- Identification of corruption of stored data is done using Recursive binary search.
- To ensure data privacy it employs the random masking technique.
- Supports multiusers and does not support the batch auditing for multi-clouds, it is possible to extend PP-PDP II.

The second technique tested on this framework uses the *Bilinearity property of the bilinear pairing*. Yang and Jia (2013) proposed the *Efficient Privacy-Preserving Public Auditing* EPP-PDP technique (Homomorphic Verifiable Tags (HVT) and bilinear pairing), which uses, an encrypted proof with the challenge stamp by using the Bilinearity property of the bilinear pairing, such that the auditor cannot decrypt it. The TPA verifies the correctness of the proof without decrypting it. This multi-cloud batch auditing protocol does not require any additional organizer and it supports the batch auditing for multiple owners. This scheme incurs less communication cost and less computation cost of the auditor by moving the computing loads of auditing from the auditor to the server, which greatly improves the auditing performance and can be applied to large scale cloud storage systems.

The third technique Sec-PDP tested on this framework uses the CBS, Designated verifier signature to offer both privacy and secuirty. Wei et al. (2014) proposed Sec-Cloud, a privacy cheating deterrence and secure computation auditing protocol to avert an adversary from accomplishing a sensitive cloud data by using designated verifier signature schemes (Huang et al. 2011a, b; Zhang and Mao 2008). Commitment-Based Sampling (CBC) technique, used in the conventional grid computing (Du et al. 2004) diminishes the computational and communication overhead; Sec-Cloud provides the batch verification for multi-users by using identity based aggregate signatures.

Succeeding the symbolization (as referred in Table 1) adapted by Shen et al. (2019), a file M is divided into n blocks, $M = \{m_1, m_2, \ldots, m_n\}$. Let P denote the prover (server), V denote the verifier (client), η denote the file's identifier, and ω denote local client state. We represent unspecified values with a, "?", symbol. The generic PDP scheme is deliberated as a five-tuple of algorithms, {KeyGen; Tag; Challenge; Proof; Verify}, each described in the subsequent section. The generic

PDP mechanism used generally comprises of four main phases—setup, challenge, proof, and verification.

6.1 Phase 01-setup phase

The client C, may or may not be the owner of the file (F), but has possession to the file. The client C, creates a public and private key pair, tags the input file, and uploads the file and tag data to storage, removing it from native storage. The input data/file F is divided into *n* blocks and the exclusive tag (metadata) for each block is calculated using the distinctive formula.

The major computation activities in the initial setup phase are user key generation, signature generation, privacy conservation and tag generation.

6.1.1 Algorithm 01: User Key-Generation algorithm (GenKey)

Upon assuming a security parameter as input, this algorithm creates a key pair (secret key, public key) as output. It is a Probabilistic key generation algorithm and it is run by client to setup the scheme, k is input and $\mathrm{s_k}$, $\mathrm{p_k}$ are outputs.

$$\mathrm{GenKey}\left(1^{\mathrm{k}}\right) \rightarrow \left(\mathrm{s_k}, \mathrm{p_k}\right)$$

The tasks accomplished by the User Key Generation algorithm are:

- User registers, by using the registration process.
- Secret key is generated for the users according to their identities (e.g. name, mail-id, contact number etc.,).
- User gets the signal/provenance to access the application once the key is received.

6.1.2 Algorithm 02: Signature—Generation algorithm (GenSig)

User creates verification metadata, and furthermore it encompasses MAC, signatures or other related information. In this algorithm, the inputs are namely, F (File),$\mathbf{s_{kID}}$ (User's private key), $\mathbf{s_{sk}}$ (user's signing private key), name (file identifier name) and outputs are $\mathbf{F}^*$(Blinded file), $\boldsymbol{\phi}$ (Corresponding signature set), τ (File tag).

$$\mathrm{SigGen}\left(\mathrm{F}, \mathbf{s_{kID}}, \mathbf{s_{sk}}, \mathrm{name}\right) \rightarrow (\mathbf{F}^*, \boldsymbol{\Phi}, \tau)$$

The tasks accomplished by the Signature—Generation algorithm are:

- user creates verification metadata, and
- furthermore it encompasses MAC, signatures or other related information.

6.1.3 Algorithm 03: Tag—Generation algorithm (GenTag)

Given a data block, using a hash function and keys as inputs, the Tag Generation algorithm gives tags as output. It is a probabilistic tag generation algorithm. It is run by the client to generate the verification metadata or verifiable tags for data. In this algorithm, b, s_k,p_k are inputs and Tb $_i$ is output.

$$\mathrm{Tb_i} \leftarrow \mathrm{GenTag}\left(\mathrm{p_k},\ \mathrm{s_k},\ \mathrm{b}_i\right)\ \text{for all } 1\ \leq\ i\ \leq\ n$$

6.1.4 Algorithm 04: Sensitive Information Sanitization algorithm (privacy conservation)

In order to preserve the personal sensitive information from the sanitizer, the user ID should blind the data blocks corresponding to the personal sensitive information of the original file F before sending it to the sanitizer. It takes as input the blinded file $\mathbf{F}^*$ and its signature set $\boldsymbol{\phi}$. It outputs the sanitized file F^{*s} and its corresponding signature set $\boldsymbol{\phi}^s$

The tasks accomplished by the Sanitization algorithm are:

- User blinds the Data blocks (enclosing the sensitive information) and produces the equivalent signatures. The signature are used to assure the validity of the file and to verify the integrity of the file.
- The blinded file and its equivalent signatures are sent to the sanitizer.
- The sanitizer sanitizes these blinded file and its equivalent signatures to generate the sanitized data blocks.
- The signatures of sanitized data blocks are transformed into valid ones for the sanitized file.

The sanitizer uploads the transformed sanitized file and its equivalent signatures to the cloud.

6.2 Phase 02-data uploading phase

Data is split and uploaded.

The tasks accomplished in the data uploading phase are:

- File is subject to the process of encryption.
- Encrypted information is generated.
- Last encrypted record is fragmented into several blocks using the process of dynamic block generation and signatures are stored in file system.
- Files are transferred and saved in the cloud server.

6.3 Phase 03-challenge phase

The client creates a challenge for a specified number of file blocks, by choosing some data blocks arbitrarily (as a challenge by using pseudo-random permutation), and sends the challenge to the prover, with the intent of auditing the cloud and to check the correctness of the stored data.

The activities accomplished in the challenge phase are:

- User/client (C) generates a challenge (chal) that, indicates the specific blocks (for which user/client (C) wants a proof of possession) are correct.
- User/client then sends challenge (chal) to server (S).
- Server (S) runs proof of possession (V) ← GenProof (p_k, F, chal, Σ) and sends to User/Client (C) the Proof of Possession (V).
- Finally, User/Client (C) can check the validity of the Proof of Possession (V) by running CheckProof (p_k, s_k, chal, V).

6.3.1 Algorithm 05: challenge—generation algorithm

The Challenge—Generation algorithm is a Probabilistic polynomial time algorithm and is run by the client. Client/user uses the Challenge Generation algorithm to produce a challenge (chal). By selecting random values, the Challenge Generation algorithm generates challenge (chal) as output, which is then sent to the prover/verifier during an audit. In this algorithm, the input is the security parameter (k) and output is a challenge (chal)

$$\text{GenChal}\ (k) \rightarrow \text{chal}$$

6.4 Phase 04-Proof Generation phase

Choose some data blocks arbitrarily as a challenge by using pseudo-random permutation, and audit the cloud to check the correctness of the stored data.

The TPA delivers a challenge or an audit message to ensure that the cloud server has retained the data file F appropriately. Upon executing GenProof, the cloud server will derive a response message from a function of the stored data file F. Using the verification metadata, the TPA verifies the response via. VerifyProof and by using the verification metadata.

6.4.1 Algorithm 06: Proof Generation algorithm (GenProof)

The Proof Generation algorithm or challenge generation is run by the server in order to create and validate a proof of possession. Prover/verifier creates a short integrity check over the customary challenge message as a proof message—that usually includes the aggregation of the blocks and the tags—and sends it to the verifier.

In this algorithm, the inputs are namely, public key (p_k), private key (s_k), Challenge (chal), Proof Of Possesion (V) and the outputs are namely, Output is correct or Proof Of data Possesion (V) for the blocks determined by Challenge (chal).

$$\text{GenProof}\ (p_k,\ F,\ \text{chal}, \Sigma) \rightarrow V$$

The tasks accomplished by the Proof Generation algorithm are:

- Client sends request to auditor to endorse the trustworthiness of the information.

Auditor completes remote data integrity checking on cloud data, by auditing the file block by block (checking one by one).

6.5 Phase 05-Proof Verification phase

The proof of possession is returned to the client, who validates the proof. Verifier authenticates the proof message relevant to the proof and challenge messages.

6.5.1 Algorithm 07: Proof Verification algorithm or check proof algorithm

Client/user uses the Proof Verification algorithm to validate the proof of possession (V). In this algorithm, inputs are namely, the public and private key pair{$\mathbf{p_k}$, $\mathbf{s_k}$}, challenge (chal).

Given generated proofs, "chal", and secret key as inputs, this algorithm verifies the proof of data possession and produces accept or reject as output. Therefore,upon effective validation it returns 1/success, else it return 0/failure.

$$\text{VerifyProof}(\text{p}_\text{k}, \text{s}_\text{k}, \text{chal}, \text{V}) \rightarrow (\text{"success"}, \text{"failure"})$$

7 Results and analysis

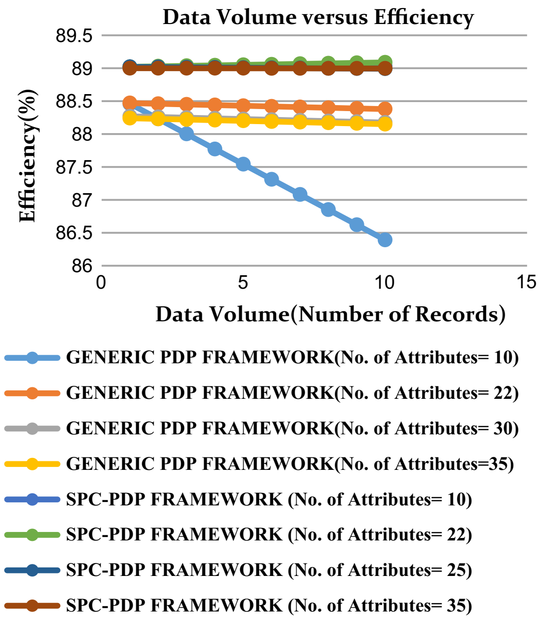
The list of appraisal parameters used for verifying and assessing the worth of the proposed SPC-PDP are given below, of which the efficiency, integrity are frequently used in our experiments.

1. *Data integrity* is a measure of the validity and fidelity of a data object. Data integrity service maintains information precisely as it was inputted, and is auditable to confirm its reliability. Data must be kept free from inaccuracies that can occur either accidentally (e.g. through programming errors), or maliciously (e.g. through breaches or hacks).
2. *Privacy* In this work for quantifying privacy it is estimated with c % confidence that a value x lies in the interval [x_1; x_2], then the interval width (x_2 - x_1) defines the amount of privacy at c % confidence level.
3. *Efficiency* To evaluate the efficiency of a searchable encryption scheme Sedghi (2012) proposes the following complexity aspects:

    - The complexity to create the searchable cipher text, the trapdoor and to perform the search (Computational Complexity).
    - The complexity to send the trapdoor and the searchable cipher text from the client to the server (Communication Complexity).
    - The complexity to store the public key, secret key, searchable cipher text and the trapdoor (Storage Complexity).

Figure 2, shows the efficiency obtained by the SPC-PDP framework, in comparison with the conventional frameworks. As the number of records increases the efficiency is more or less sustained in both the earlier and proposed frameworks. The

Fig. 2 Data volume (number of records) versus efficiency



number of attributes in a record is kept a constant and the number of records are gradually increased to gauge the efficiency of the framework. It is observed that as we increase the number of records the SPC-PDP framework more efficiently and proficiently achieves secure auditing, privacy, integrity and outclasses the erstwhile generic PDP framework.

Fig. 3 Attributes versus efficiency

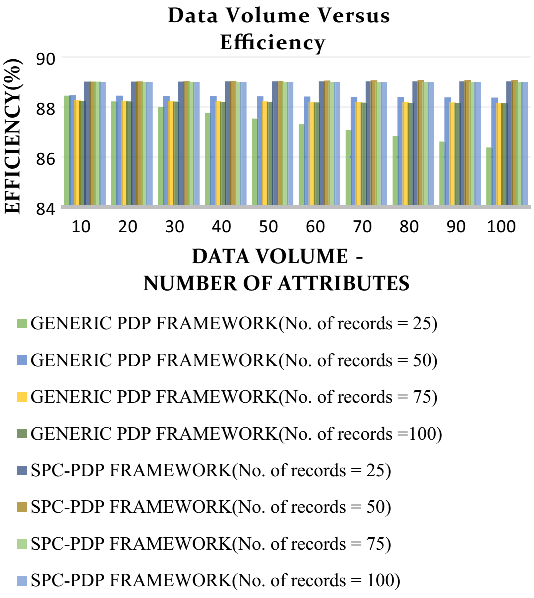


Figure 3, shows the efficiency obtained by the SPC-PDP framework, in comparison with the conventional frameworks. As the number of attributes increases the efficiency is more or less sustained in both the earlier and proposed frameworks. The number of records is kept a constant and the number of attributes are gradually increased to gauge the efficiency of the framework. It is observed that as we increase the number of attributes the SPC-PDP framework more efficiently achieves secure auditing, privacy, and integrity and outclasses the erstwhile generic PDP framework.

The running time increases with rise in data volume and is more or less the same for both the frameworks—SPC-PDP framework, Generic PDP framework (as seen in Fig. 4). The amount of time CSP takes to retrieve the user data-running time, is just very few seconds. After the execution of this work the time taken for the challenge and response was noted. From that, it is inferred that as the file size increases, time taken also increases. Time duration of different file sizes varies only in few milli seconds. Since, the time duration varies in few milli seconds the user/verifier cannot perceive the difference in challenge and response.

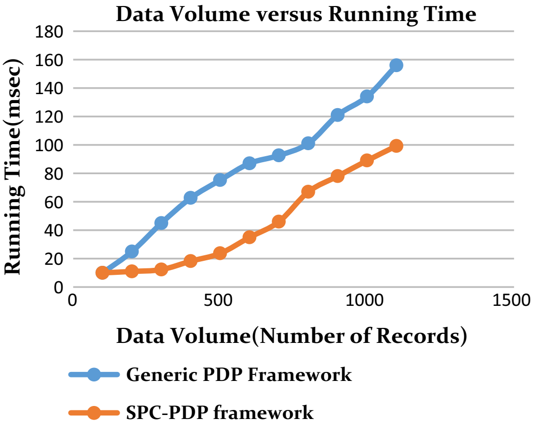
Figure 5, shows the reliability obtained by the SPC-PDP framework, in comparison with the conventional frameworks. As the number of files increases, the reliability is more or less persistent in both the earlier and proposed frameworks.

Abridging the above key points along with other features, a perfect equilibrium is thus gifted leading to an increase in data utility, increase in privacy and integrity of cloud data storage. As we use different random values for each data item, the degree of privacy is high and almost near to 100%. Based on the above factor we are going to analyze the SPC-PDP approach that we have implemented.

The performance of the various algorithms are tested on the proposed SPC-PDP framework is measured based on the ***time*** necessary by each algorithm to secure and privacy conserve the data.

Figures 6, 7 and 8 shows the time taken for the deployment of the techniques of three PDP schemes, on the SPC-PDP framework for a fixed file size of 100 MB with different block sizes. The techniques of three PDP schemes (PP-PDP, EPP-PDP, Sec-PDP), have been implemented, in the generic PDP framework and the

Fig. 4 Data volume versus running time



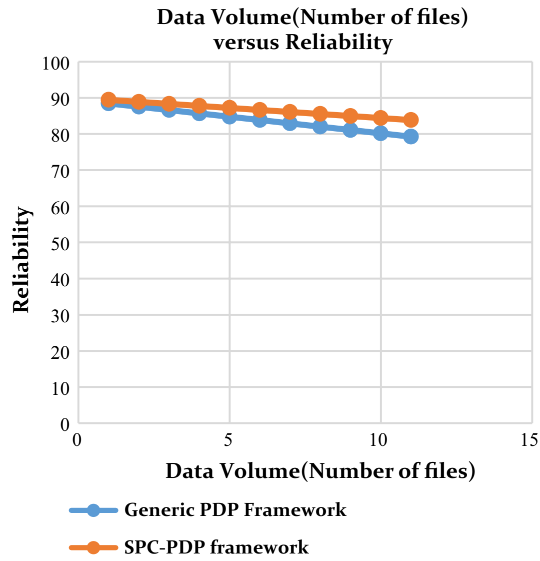


Fig. 5 Data volume versus reliability

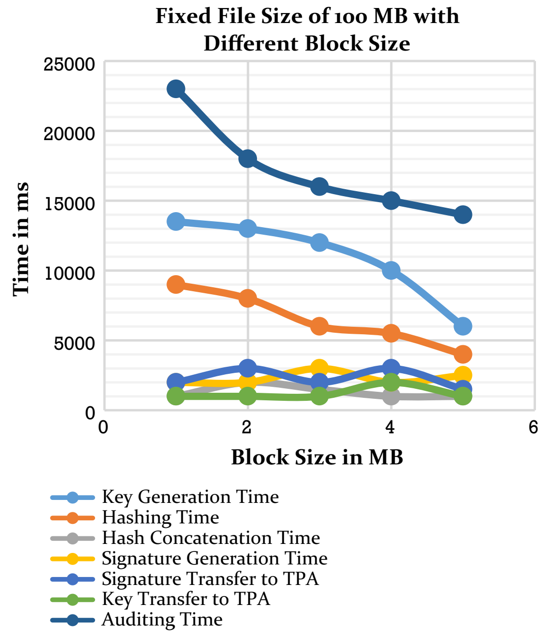


Fig. 6 Time consumed for PP-PDP

proposed SPC-PDP framework. The time taken for the various activities of three PDP schemes (PP-PDP, EPP-PDP, Sec-PDP) were noted. The time taken for key generation, hash concatenation, signature transfer to TPA, auditing, hashing, signature generation and key transfer to TPA have been taken into consideration for analysis. It is established that time consumed of the sophisticated PDP schemes (PP-PDP,

Fig. 7 Time consumed for EPP-PDP

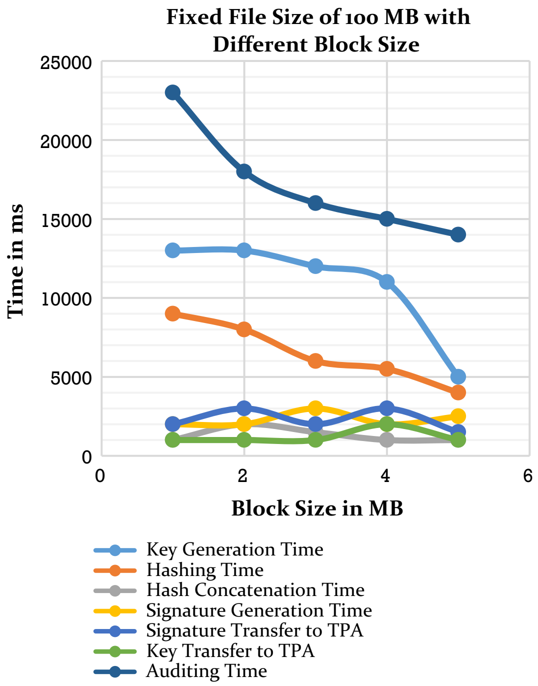
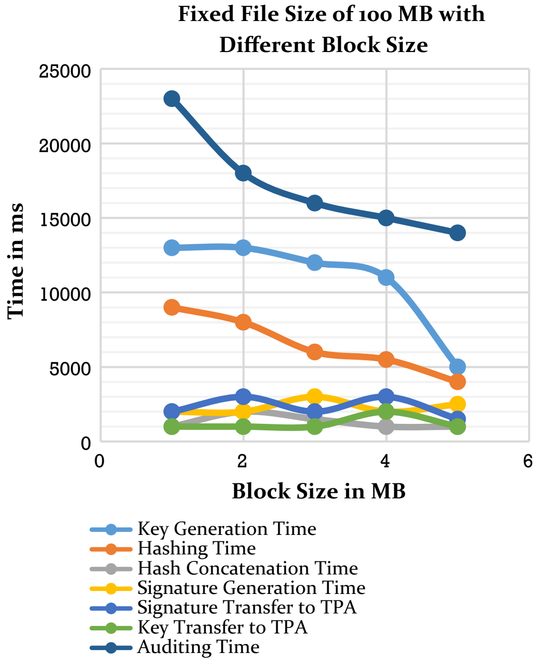


Fig. 8 Time consumed for Sec-PDP



EPP-PDP, Sec-PDP) are nearly identical to those implemented on the simple PDP frameworks; however, tag/preprocessing in schemes utilizing asymmetric key operations have a noteworthy control on the total time differences among the schemes. It is thus indomitable that the total time consumed of schemes utilizing the symmetric key primitives are analogous, whereas the time of schemes utilizing public key primitives are more costly compared to the other schemes at large file sizes.

Further, the Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework, deployment results recorded confirm that the SPC-PDP framework can adeptly accomplish secure auditing and outperform the erstwhile ones, in terms of, efficiency, reliability, cost-effective computational time, communication time. The SPC-PDP architecture is thus proved to be, a promising solution to the challenges soaring due to the state-of-the-art improvements in digital technology.

8 Conclusions

In this paper, the integrity and privacy solutions that are mandatory to fix the problems arising due to genesis of cloud computing technologies are analyzed. A comprehensive appraisal of the correlated up-to-date research work on the prevailing Privacy Preserving Provable Data Possession (PDP) mechanism, its evolution, used for conserving privacy and integrity of the shared data in storage applications; yearns for a public auditing system of health care data storage security and a privacy-conserving generic framework, which can weather the evolving digital storage technologies. The Privacy conserving Provable Data Possession framework should aid an external auditor to audit user's outsourced health care data in the cloud deprived of the wisdom of the health care data content. The SPC-PDP framework is a scalable, efficient public auditing system, which can accomplish batch auditing (where multiple delegated auditing tasks from different users can be performed simultaneously by the TPA). The Secure Privacy Conserving Provable Data Possession (SPC-PDP) framework, deployed shows that it can competently accomplish secure auditing and privacy; and outclass the erstwhile ones, in terms of scalability, reliability, efficiency, privacy, and computation time. The SPC-PDP framework is no doubt, a promising solution to the challenges soaring due to the state-of-the-art improvements in digital technology. Moreover, the deployment of SPC-PDP does not entail the server to offer any extra services or APIs. Since, SPC-PDP does not entail server-side computation, it is surmised that the SPC-PDP scheme is highly practical to be deployed. The *communication cost* incurred through the barter of information among a number of collaborating sites, should be painstaking. It is crucial that this cost must be reserved to a minimum for a distributed Privacy Preserving Data Mining algorithm. Any future PDP schemes should be technologically advanced with the realistic cloud restrictions in mind.

9 Future work

The proposed framework can be implemented with other data integrity practices like, proof of ownership (POR), third party auditing methods—(MAC, signature based, and MD5 based), and Encryption algorithms, used for conserving privacy and integrity of the shared data in storage applications. The future direction for research will be to deal with the heterogeneous data, such as audio, video, image, or text message which, have diverse provisioning necessities that must be consistent to offer coherent knowledge to the cloud client. The next promising direction of research will be to proactively, handle the risk like uncertainty, unpredictability of the happenings in the communication path between the server and TPA, and study the framework of an interaction-based system using a graphical dynamic system. Another promising direction is to Commission a separate architecture (that is needed from the data auditing perspective), and to tackle the technical challenges of auditing services. There are several sources like devices with diverse backhaul networks, such as 2G, 3G, LTE, and 4G; that generate and transfer data to the cloud for storage. These multiple devices not only have different architectures, but also have different network delivery systems and therefore these devices must be synchronized in order to offer seamless connections.

References

Abbas A, Khan SU (2014) A review on the state-of-the-art privacy-preserving approaches in the e-health clouds. IEEE J Biomed Health Inf 18(4):1431–1441

Abbas A, Khan M, Ali M, Khan S, Yang L (2015) A cloud based framework for identification of influential health experts from Twitter. In: Proceedings of the 15th international conference on scalable computing and communications (ScalCom)

Ahuja S, Mani S, Zambrano J (2012) A survey of the state of cloud computing in healthcare. Netw Commun Technol 1:12–19

Akinyele J, Lehmann C, Green M, Pagano M, Peterson Z, Rubin A (2010) Self-protecting electronic medical records using attribute-based encryption. In: Technical report 2010/565, Cryptology e-Print Archive

Ateniese G, Di Pietro R, Mancini LV, Tsudik G (2008) Scalable and efficient provable data possession. In: Proceedings of the 4th international conference on security and privacy in communication networks, p 9. ACM

Ateniese G, Kamara S, Katz J (2009) Proofs of storage from homomorphic identification protocols. In: International conference on the theory and application of cryptology and information security, pp 319–333. Springer, Berlin

Bhadauria R, Chaki R, Chaki N, Sanyal S (2011) A survey on security issues in cloud computing. arXiv preprint arXiv:1109.5388, pp 1–15

Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I (2009) Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Future Gener Comput Syst 25(6):599–616

Chen Y, Lu J, Jan J (2012) A secure EHR system based on hybrid clouds. J Med Syst 36(5):3375–3384

Du W, Jia J, Mangal M, Murugesan M (2004) Uncheatable grid computing. In: 24th international conference on distributed computing systems, proceedings, pp 4–11. IEEE

Erway CC, Küpçü A, Papamanthou C, Tamassia R (2015) Dynamic provable data possession. ACM Trans Inf Syst Secur 17(4):15

Fu A, Yu S, Zhang Y, Wang H, Huang C (2017) NPP: a new privacy-aware public auditing scheme for cloud data sharing with group users. IEEE Trans Big Data. https://doi.org/10.1109/TBDATA.2017.2701347

Gasti P, Ateniese G, Blanton M (2010) Deniable cloud storage: sharing files via public-key deniability. In: Proceedings of the 9th annual ACM workshop on privacy in the electronic society, pp 31–42. ACM

Gitanjali J, Banu SN, Geetha Mary A, Indumathi J, Uma GV (2007) An agent based burgeoning framework for privacy preserving information harvesting systems. Int J Comput Sc Netw Secur 7(11):268–276

Gitanjali J, Banu SN, Indumathi J, Uma GV (2008) A panglossian solitary-skim sanitization for privacy preserving data archaeology. Int J Electr Power Eng 2(3):154–165

Gitanjali J, Ghalib MdR, Murugesan K, Indumathi J, Manjula D (2009a) An object-oriented scaffold premeditated for privacy preserving data mining of outsourced medical data. Int J Softw Eng Appl (**in press**)

Gitanjali J, Ghalib MdR, Murugesan K, Indumathi J, Manjula D (2009b) A hybrid scheme Of data camouflaging for privacy preserved electronic copyright publishing using cryptography and watermarking technologies. Int J Secur Appl

Grobauer B, Walloschek T, Stocker E (2011) Understanding cloud computing vulnerabilities. IEEE Secur Priv 9(2):50–57

Guan C, Ren K, Zhang F, Kerschbaum F, Yu J (2015) Symmetric-key based proofs of retrievability supporting public verification. In: European symposium on research in computer security (pp 203–223). Springer, Cham

Haas S, Wohlgemuth S, Echizen I, Sonehara N, Muller G (2011) Aspects of privacy for electronic health records. Int J Med Inform 80(2):e26–e31

He J, Zhang Z, Li M, Zhu L, Hu J (2018) Provable data integrity of cloud storage service with enhanced security in the internet of things. IEEE Access 7:6226–6239

Huang D, Misra S, Verma M, Xue G (2011a) PACP: an efficient pseudonymous authentication-based conditional privacy protocol for VANETs. IEEE Trans Intell Transp Syst 12(3):736–746

Huang Q, Yang G, Wong DS, Susilo W (2011b) Efficient strong designated verifier signature schemes without random oracle or with non-delegatability. Int J Inf Secur 10(6):373–385

Indumathi J (2012) A generic scaffold housing the innovative modus operandi for selection of the superlative anonymisation technique for optimized privacy preserving data mining. In: Karahoca A (ed) Data mining applications in engineering and medicine, Chapter 6. InTech, London, pp 133–156. ISBN 9535107200 9789535107200

Indumathi J (2013a) Amelioration of anonymity modus operandi for privacy preserving data publishing. In: Amine A, Mohamed OA, Benatallah B (eds) Network security technologies: design and applications, chapter 7. IGI Global, Hershey, pp 96–107

Indumathi J (2013b) An enhanced secure agent-oriented burgeoning integrated home tele health care framework for the silver generation. Int J Adv Network Appl 04(04):16–21. Special issue on "Computational intelligence—a research perspective" held on "21st–22nd Feburary 2013"

Indumathi J (2013c) State-of-the-art in reconstruction-based modus operandi for privacy preserving data dredging. Int J Adv Netw Appl 04(04):9–15. Special issue on "Computational intelligence—a research perspective" held on "21st–22nd Feburary 2013"

Indumathi J, Uma GV (2007a) Customized privacy preservation using unknowns to stymie unearthing of association rules. J Comput Sci 3(12):874–881

Indumathi J, Uma GV (2007b) Using privacy preserving techniques to accomplish a secure accord. Int J Comput Sci Netw Secur 7(8):258–266

Indumathi J, Uma GV (2008a) A bespoked secure framework for an ontology-based data-extraction system. J Softw Eng 2(2):1–13

Indumathi J, Uma GV (2008b) A new flustering approach for privacy preserving data fishing in telehealth care systems. Int J Healthc Technol Manag 9(5–6):495–516 (**Special issue on: "Tele-healthcare system implementation, challenges and issues"**)

Indumathi J, Uma GV (2008c) A novel framework for optimized privacy preserving data mining using the innovative desultory technique. Int J Comput Appl Technol 35(2/3/4):194–203 (**Special Issue on: "Computer applications in knowledge-based systems"**)

Indumathi J, Uma GV (2008d) An aggrandized framework for genetic privacy preserving pattern analysis using cryptography and contravening—conscious knowledge management systems. Int J Mol Med Adv Sci 4(1):33–40

Jain P, Gyanchandani M, Khare N (2016) Big data privacy: a technological perspective and review. J Big Data 3(1):25

Joshi B, Vijayan AS, Joshi BK (2011) Securing cloud computing environment against DDoS attacks. IEEE, pp 1–5

Kaletsch A, Sunyaev A (2011) Privacy engineering: personal health records in cloud computing environments. In: Proceedings of the 32nd international conference on information systems (ICIS), pp 1–11

Khedr WI, Khater HM, Mohamed ER (2019) Cryptographic accumulator-based scheme for critical data integrity verification in cloud storage. IEEE Access 7:65635–65651

Kuo AH (2011) Opportunities and challenges of cloud computing to improve health care services. J Med Internet Res 13(3):e67. https://doi.org/10.2196/jmir.1867

Juels A, Kaliski Jr, BS (2007) PORs: Proofs of retrievability for large files. In: Proceedings of the 14th ACM conference on computer and communications security, pp 584–597. ACM

Li J (2013) Electronic personal health records and the question of privacy. Computer. https://doi.org/10.1109/mc.2013.225

Li M, Yu S, Ren K, Lou W (2010) Securing personal health records in cloud computing: patient-centric and fine-grained data access control in multi-owner settings. International conference on security and privacy in communication networks. Springer, Berlin, Heidelberg, pp 89–106

Lin H, Shao J, Zhang C, Fang Y (2013) CAM: cloud-assisted privacy preserving mobile health monitoring. IEEE Trans Inf Forensics Secur 8(6):985–997

Liu W (2012) Research on cloud computing security problem and strategy. IEEE, pp 1216–1219

Lounis A, Hadjidj A, Bouabdallah A, Challal Y (2015) Healing on the cloud: secure cloud architecture for medical wireless sensor networks. Future Gener Comput Syst. https://doi.org/10.1016/j.future.2015.01.009

Luo Y, Xu M, Fu S, Wang D, Deng J (2015) Efficient integrity auditing for shared data in the cloud with secure user revocation. In 2015 IEEE Trustcom/BigDataSE/ISPA; vol 1, pp 434–442. IEEE

Mandagere N, Zhou P, Smith MA, Uttamchandani S (2008) Demystifying data deduplication. In: Proceedings of the ACM/IFIP/USENIX Middleware'08 conference companion, pp 12–17. ACM

Mashima D, Ahamad M (2012) Enhancing accountability of electronic health record usage via patient-centric monitoring. In: Proceedings of the 2nd ACM SIGHIT international health informatics symposium, IHI'12, pp 409–418. ACM

Metri P, Sarote G (2011) Privacy issues and challenges in cloud computing. Int J Adv Eng Sci Technol 5:1–6

Meyer DT, Bolosky WJ (2012) A study of practical deduplication. ACM Trans Storage 7(4):14. https://doi.org/10.1145/2078861.2078864

Murugesan K, Gitanjali J, Indumathi J, Manjula D (2009) Sprouting modus operandi for selection of the best PPDM technique for health care domain. Int J Conf Recent Trends Comput Sci 1(1):627–629

Murugesan K, Indumathi J, Manjula D (2010a) An optimized intellectual agent based secure decision system for health care. Int J Eng Sci Technol 2(5):3662–3675

Murugesan K, Indumathi J, Manjula D (2010b) A framework for an ontology-based data-gleaning and agent based intelligent decision support PPDM system employing generalization technique for health care. Int J Comput Sci Engi 2(5):1588–1596

Narayanan A, Huey J, Felten EW (2016) A precautionary approach to big data privacy. Data protection on the move. Springer, Dordrecht, pp 357–385

Prakash D, Murugesan K, Indumathi J, Manjula D (2009) A novel cardiac attack prediction and classification using supervised agent techniques. CiiT Int J Artif Intell Syst Mach Learn 1(2):59

Ren K, Wang C, Wang Q (2012) Security challenges for the public cloud. IEEE Internet Comput 16(1):69–73

Ren Y, Xu J, Wang J, Kim JU (2013) Designated-verifier provable data possession in public cloud storage. Int J Secur Appl 7(6):11–20

Rong C, Nguyen ST, Jaatun MG (2013) Beyond lightning: a survey on security challenges in cloud computing. Comput Electr Eng 39(1):47–54

Satheesh Kumar K, Indumathi J, Uma GV (2008) Design of smoke screening techniques for data surreptitiousness in privacy preserving data snooping using object oriented approach and UML. IJCSNS Int J Comput Sci Netw Secur 8(4):106–115

Sebé F, Domingo-Ferrer J, Martinez-Balleste A, Deswarte Y, Quisquater JJ (2008) Efficient remote data possession checking in critical information infrastructures. IEEE Trans Knowl Data Eng 20(8):1034–1038

Sedayao J (2012) Enhancing cloud security using data anonymization. White Paper, Intel Coporation

Sedghi S (2012) Towards provably secure efficiently searchable encryption. University of Twente, Enschede

Shen W, Yang G, Yu J, Zhang H, Kong F, Hao R (2017) Remote data possession checking with privacy-preserving authenticators for cloud storage. Future Gener Comput Syst 76:136–145

Shen W, Qin J, Yu J, Hao R, Hu J (2019) Enabling identity-based integrity auditing and data sharing with sensitive information hiding for secure cloud storage. IEEE Trans Inf Forensics Secur 14(2):331–346

Sookhak M, Gani A, Talebain H, Akhunzada A, Khan S, Buyya R, Zomaya A (2015) Remote data auditing in cloud computing environments: a survey, taxonomy, and open issues. ACM Comput Surv 47(4):65:1–65:34

Tang J, Cui Y, Li Q, Ren K, Liu J, Buyya R (2016) Ensuring security and privacy preservation for cloud data services. ACM Comput Surv 49(1):13:1–13:39

Tian H, Chen Y, Chang CC, Jiang H, Huang Y, Chen Y, Liu J (2017) Dynamic-hash-table based public auditing for secure cloud storage. IEEE Trans Serv Comput 10(5):701–714

Vasudevan V, Sivaraman N, SenthilKumar S, Muthuraj R, Indumathi J, Uma GV (2007) A comparative study of SPKI/SDSI and K-SPKI/SDSI systems. Inf Technol J 6(8):1208–1216

Wang Q, Wang C, Li J, Ren K, Lou W (2009) Enabling public verifiability and data dynamics for storage security in cloud computing. In: European symposium on research in computer security. Springer, Berlin, pp 355–370

Wang C, Wang Q, Ren K, Lou W (2010) Privacy-preserving public auditing for data storage security in cloud computing. In: 2010 proceedings IEEE infocom. IEEE, pp 1–9

Wang Q, Wang C, Ren K, Lou W, Li J (2011) Enabling public auditability and data dynamics for storage security in cloud computing. IEEE Trans Parallel Distrib Syst 22(5):847–859

Wang B, Li B, Li H (2012) Knox: privacy-preserving auditing for shared data with large groups in the cloud. In: International conference on applied cryptography and network security. Springer, Berlin, pp 507–525

Wang C, Chow SS, Wang Q, Ren K, Lou W (2013) Privacy-preserving public auditing for secure cloud storage. IEEE Trans Comput 62(2):362–375

Wang B, Li B, Li H (2014a) Oruta: privacy-preserving public auditing for shared data in the cloud. IEEE Trans Cloud Comput 2(1):43–56

Wang B, Li B, Li H (2014b) Oruta: privacy-preserving public auditing for shared data in the cloud. IEEE Trans Cloud Comput 2(1):43–56

Wang B, Li B, Li H (2015) Panda: public auditing for shared data with efficient user revocation in the cloud. IEEE Trans Serv Comput 8(1):92–106

Wang H, He D, Tang S (2016a) Identity-based proxy-oriented data uploading and remote data integrity checking in public cloud. IEEE Trans Inf Forensics Secur 11(6):1165–1176

Wang H, He D, Yu J, Wang Z (2016b) Incentive and unconditionally anonymous identity-based public provable dat possession. IEEE Trans Serv Comput. https://doi.org/10.1109/TSC.216.2633260

Wei L, Zhu H, Cao Z, Dong X, Jia W, Chen Y, Vasilakos AV (2014) Security and privacy for storage and computation in cloud computing. Inf Sci 258:371–386

Worku SG, Ting Z, Zhi-Guang Q (2012) Survey on cloud data integrity proof techniques. In: 2012 Seventh Asia joint conference on information security. IEEE, pp 85–91

Worku SG, Xu C, Zhao J, He X (2014) Secure and efficient privacy-preserving public auditing scheme for cloud storage. Comput Electr Eng 40(5):1703–1713

Wu R, Ahn GJ, Hu H (2012) Secure sharing of electronic health records in clouds. In: Proceedings of the 8th international conference on collaborative computing: networking, applications and worksharing (CollaborateCom), pp 711–718

Xiao Z, Xiao Y (2012) Security and privacy in cloud computing. IEEE Commun Surv Tutorials 15(2):843–859

Yan H, Li J, Zhang Y (2019) Remote data checking with a designated verifier in cloud storage. IEEE Syst J. https://doi.org/10.1109/JSYST.2019.2918022

Yang K, Jia X (2012) Data storage auditing service in cloud computing: challenges, methods and opportunities. World Wide Web 15(4):409–428

Yang K, Jia X (2013) An efficient and secure dynamic auditing protocol for data storage in cloud computing. IEEE Trans Parallel Distrib Syst 24(9):1717–1726

Yang JJ, Li JQ, Niu Y (2015) A hybrid solution for privacy preserving medical data sharing in the cloud environment. Future Gener Comput Syst 43–44:74–86

Yang G, Yu J, Shen W, Su Q, Fu Z, Hao R (2016) Enabling public auditing for shared data in cloud storage supporting identity privacy and traceability. J Syst Softw 113:130–139

Yu J, Wang H (2017) Strong key-exposure resilient auditing for secure cloud storage. IEEE Trans Inf Forensics Secur 12(8):1931–1940

Yu J, Ren K, Wang C, Varadharajan V (2015) Enabling cloud storage auditing with key-exposure resistance. IEEE Trans Inf Forensics Secur 10(6):1167–1179

Yu J, Ren K, Wang C (2016) Enabling cloud storage auditing with verifiable outsourcing of key updates. IEEE Trans Inf Forensics Secur 11(6):1362–1375

Yu Y, Au MH, Ateniese G, Huang X, Susilo W, Dai Y, Min G (2017) Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage. IEEE Trans Inf Forensics Secur 12(4):767–778

Zhang R, Liu L (2010) Security models and requirements for healthcare application clouds. In: Proceedings of the 3rd IEEE international conference on cloud computing (CLOUD), pp 268–275

Zhang J, Mao J (2008) A novel ID-based designated verifier signature scheme. Inf Sci 178(3):766–773

Zhang Y, Yu J, Hao R, Wang C, Ren K (2018) Enabling efficient user revocation in identity-based cloud storage auditing for shared big data. IEEE Trans Dependable Secure Comput. https://doi.org/10.1109/TDSC.2018.2829880

Zhu Y, Hu H, Ahn GJ, Yu M (2012) Cooperative provable data possession for integrity verification in multicloud storage. IEEE Trans Parallel Distrib Syst 23(12):2231–2244

Zissis D, Lekkas D (2012) Addressing cloud computing security issues. Future Gener Comput Syst 28(3):583–592

Zyskind G, Nathan O, Pentland A (2015) Enigma: Decentralized computation platform with guaranteed privacy. *arXiv preprint* arXiv:1506.03471

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Indumathi Jayaraman[1] · **Mokhtar Mohammed**[1]

Mokhtar Mohammed
mokhtar201328@gmail.com

[1] Department of Information Science and Technology, Anna University, Chennai,
Tamilnadu 600 025, India