A survey on data integrity verification schemes using blockchain technology in Cloud Computing Environment

K GANGADEVI ¹ and R RENUGA DEVI ²

¹² Department of Computer Science, VELS University, Pallavaram, Chennai, India
¹seenuganga@yahoo.co.in and ²nicrdevi@gmail.com

Abstract. Cloud computing is the delivery of different services for business or any organization to store their data and utilize the various resources over the internet. The resources of cloud computing is storage, software, computing hardware and services. Usefulness of the cloud computing infrastructure is to grant services based on demand to the user. The security is a major concern in the cloud computing that are many modes of securities including accessibility, authentication, data integrity, data recovery and confidentiality. The data integrity is major role in the cloud computing security and storage service. Hence, the data integrity scheme provides the proof the truthfulness of the stored data in the cloud. In this paper, we are conducting a brief survey on data integrity verification schemes for cloud computing based on blockchain technology. The aim of this survey is to enhance the understanding the security issues and discuss the importance of data integrity scheme for outsourced data.

Keywords: cloud computing, data integrity, proof data possessions, proof of retrievability, blockchain technology, validation.

1. Introductions

The cloud computing is gained tremendous popularity due to outsourcing computation and providing storage requirements based on the user availability. The user can use the resources over the internet such as storages, services, applications, and bandwidth based on distributed shared services. The resource provider is called the cloud service provider (CSP). The cloud computing is organized into three categories based on the services. (i) Infrastructure-as-a-Service (IaaS), (ii) Platform-as-a-Service (PaaS) and (iii) Software-as-a-Service (SaaS). Three deployment models in cloud environment are private, public and Multi or Hybrid models [1]. One primary feature of cloud computing is the outsourcing the user data in cloud environment with centralized manner. The services of outsourced storage in cloud environment have an innovative yield development point by providing a comparable high scalable, low-cost, and location-independent environment for the control the data. Many cloud based storages are available today such as One Drive, Amazon, Google Drive and Drop Box.
A cloud storage service (CSS) is a critical cloud computing service that belongs to IaaS. The CSS is to mitigate the load of storage management and preservation of local storage. However, this kind of significant service affects by malicious attacks or failures both from outside and inside the cloud.

Modification, corruption and deletions. This may be caused by careless scheme preserve and disturb the cost saving behaviour. Hence, CSP required allowing an efficient auditing service to validate the integrity and accessibility of the stored data. The CSP is enclosed by a service level agreement (SLA) to make sure the security of the outsourced data [2]. The architecture of cloud computing environment is shown in Figure 1.1.

The regular solution for making sure the integrity of outsourced data is to utilize conventional cryptographic models based on signature and hashing methods. But, discussed conventional methods need a local copy of data to make sure the integrity of outsourced data [3]. This would be reached an expensive I/O cost and massive communication overhead across the networks.

In the recent years, the verifying the outsourced data integrity in cloud storage is more attention for industry persons and researchers. The aim of this paper is conducting a brief survey on data integrity scheme based on blockchain technology. First, an introduction and fundamental details of data integrity scheme, fundamental details of blockchain technology, various data integrity scheme for verifying outsourced data in cloud based on blockchain technology schemes. Finally, present the various challenges of blockchain technology when apply to data integrity scheme.

1.1 Data integrity

Data integrity is the guarantee that the digital information is uncorrupted and can be accessed or modified by approved users. The data integrity involves preserving the stability, accurateness and reliability of data. The integrity verification scheme is a significantly essential for data assurance and digital forensics. The data integrity verification process is a major task in cloud computing because of the data storage is location-independent behaviour [4]. In the earlier stage, the signature scheme and hash functions are used to check data integrity. This both scheme is want to download complete data from the cloud storage then hash value or signature value of download data are compared with local data. So, the hash tree is resolving the above mentioned fault, but it only validate the exactness of the hash tree corresponding to the cloud storage rather than the data itself. Hence, the designing the significant mechanism is to verify the data integrity significantly without having to recover the complete data.
To solve abovementioned shortcomings, many research works have been conducted on allowing public auditing scheme by employing data integrity verification methods. Generally, the data integrity is an interactive protocol running between CSSs and user. The data integrity scheme performed between CSS and user is called private auditing. In the private auditing scheme, the users required to calculate some authentication data internally and often communicate with a CSS during storing the data, which takes more computation burden.

![Architecture of data integrity scheme](image)

**Figure 1.2:** Architecture of data integrity scheme

The third-party auditor (TPA) is launch to communicate with user for auditing. The auditor can be either TPA or authentic user is called public auditing. The TPA is used to improve auditing effectiveness and reduce the computation cost. Consequently, developing a proper data integrity verification model for public auditing is considered as a vital requirement [5].

Two major types of data integrity schemes are used for public auditability of the outsourced data, such as Provable Data Possession (PDP) [6] and Proof of Retrievability (PoR) [7]. PDP scheme was introduced by Ateniese et al. (2007) [6]. In the PDP scheme, the client is sent the pre-processed the data to an untrusted server for storage. Also, it is keeping a small amount of meta-data. Later, the client requests the server to demonstrate that the stored information has not been corrupt with or deleted without downloading the authentic data. However, if the data once corrupted or deleted by the malicious user, the PDP scheme is not retrieved properly. This behavior of PDP has many shortcomings including data loss, financial scratch, and loss of trust. PoP is another data integrity scheme similar to that of PDP [7]. It has another advantage contrast with PDP that will recover the loosed data or corrupted data by using error correcting code (ECC) method.

The data integrity can be classified into two categories either deterministic or probabilistic. The deterministic scheme is necessary to access the entire file in order to find out the integrity [8]. This scheme is not suitable for large files because it will take more time for completing the process of integrity verification. The probabilistic scheme is dynamically accessing the data block to verify the data integrity [1]. This scheme is more suitable for large files to complete with less computational time. The performance of this scheme is poor when size of data is low. The nature of data is significant role in the data integrity verification process. The data can be classified based on its behavior either static nature or dynamic nature. The static nature of the data cannot modify once stored in the cloud. Static nature of the data is having small storage overhead and given that higher acceptance speed
for errors [9]. In term of dynamic nature, the data can be insertion, updation, deletion and modification in the cloud storage [10].

1.2 Architecture of data integrity scheme
The architecture of the data integrity scheme for outsourced data is discussed in this section and architecture of data integrity scheme is portrayed in Figure 2. The cloud data storage services consist of four phases which is explained as follows,

- **Data Owner (DO):** Who is utilizes the CSS for storing their huge amount of data
- **Cloud Service Provider (CSP):** who is providing the CSS and computational resources
- **Third Party Auditor (TPA):** she/he has capabilities to supervise the outsourced data under the allocation of data

The DO generally adopts a TPA to accomplish the data integrity verification procedure to evaluate the integrity of outsourced data. DO can’t be entirely authenticated the consistency of the TPA during the integrity verification process. Hence, furnish over the verification of data integrity to the malicious TPA may lead to threats of data security. In the earlier stage, the data integrity scheme is performed on the static data and then extended to dynamic data. Similarly, the private data integrity verification scheme is conducted in the earlier stage and then public data integrity scheme is launched to verify the data with the help of TPA or without.

2. Blockchain technology
Recently, cryptocurrency has concerned widespread attention from both academic and industries. The bitcoin is a first cryptocurrency and the core mechanism of blockchain technology. The blockchain is introduced by Nakamoto (2008) [11]. Blockchain is decentralized digital public ledger for distributed storage and peer-to-peer communication between nodes. Blockchain is facilitated by combining with numerous core technologies including digital signature, cryptographic hash, time stamp and distributed consensus algorithm. In blockchain, all committed transactions are stored in the form of chain of blocks. Chain of blocks is continuously increased one by one when the new blocks are created [12]. The blockchain technology is more efficient and cost effective because of transactions can store in the decentralized manner. The blockchain can be applied various fields of application including cloud computing [13], big data environment [14], financial service [15] and Internet of Things (IoT) [16].

2.1 Architecture of blockchain technology

![Architecture of Blockchain Technology](image-url)

**Figure 1.3:** Architecture of Blockchain Technology
The blockchain is an order of blocks that holds an entire list of transaction records. The architecture of blockchain shows in Figure 1.3. Each block is a directly point to the previous block using hash value that the prior block is called parent block. The first block of blockchain is called genesis block and it has no parent block. Each block of blockchain has two major components including a block header and a block body. The block header contains much information such as block version, parent hash values, Markel root hash, time stamp, nBits and nonce [17]. The block body contains a detailed transaction data and a transaction counter.

The maximum numbers of transactions are determined based on the block size and transaction. The blockchain uses method of an asymmetric cryptography to validate the transactions. A digital signature based on asymmetric cryptography is used in an unreliable environment. Each user holds either a private key or a public key. The use of private key is sign the transactions. The signed transactions are extended during the entire network that accessed by public key. The digital signature is classified into two phases including signing and verification phases. In the blockchain architecture, the elliptic curve digital signature algorithm (ECDSA) is used as a classic signature [18].

2.2 Types of blockchain technology

Three types of the blockchain technologies are private, public and consortium blockchain [19] which is represented in Figure 4.

2.2.1 Private blockchain

The private blockchain is a distributed database. This approach favours (i) by allowing few organizations with privacy requirements to implement blockchain; (ii) cryptographic audit adds value. The permissions are centralized to an organization. It is an easier method due to the small number of participating users. Hence, the transaction and processing time is low. It is an inverse of the public blockchain due to environment friendly as less computational power is required. Comparing to the public blockchain, it is less secure.

2.2.2 Public blockchain

The public blockchain allows any user to add, download data or block in the blockchain network. The public blockchain is a release for all kinds of entities to contribute in the networks. It is a completely transparent and decentralized database of the transaction like an open network. It is highly secured using cryptography and consensus protocol which is an answer for those who criticize that it’s not secured since the public blockchain is an open source and low privacy. It is less eco-friendly due to high computational power and energy is required.
Figure 2.1: Types of blockchain technology

2.2.3 Consortium blockchain
The consortium blockchain is “partially private” works across different organizations with pre-defined set of users. It is a fast, delivery with better scalability and privacy.

| Algorithms   | Energy Consumptions | Transactions/Sec | Computation cost | Example     |
|--------------|---------------------|------------------|------------------|-------------|
| PBFT         | High                | -                | Low              | Hyperledger |
| PoW          | High                | 7-30             | High             | Bitcoin     |
| PoS          | Low                 | 30-173           | Low              | Peercoin    |
| DPoS         | Very low            | 25-2000          | Low              | Bitshares   |
| Ripple       | High                | -                | Low              | -           |
| Tendermint   | Low                 | -                | Low              |             |

Table 1: Comparison of consensus algorithms

2.3 Consensus algorithms
how to achieve consensus mechanism among the undependable nodes is a transformation of the problem of Byzantine Generals (BG) [20]. In blockchain like a BG problem, there is no middle node. Nodes are no need to trust each other. Hence, some set of rules are required to make sure that the ledgers in different nodes are reliable [21]. Consensus algorithm is a mechanism to make sure the precision and reliability of information stored by all nodes in a distributed ledger. Consensus mechanism is used to authenticate every single block and eventually used to achieve the required consensus. Also, it is a process of removing invalid or contradictory transactions. The many consensus algorithms are used in blockchain technologies which are discussed in the following sub paragraphs. Table 1 illustrate the comparisons of the consensus algorithms.

2.3.1 Practical byzantine fault tolerance (PBFT)
PBF is a replication algorithm which can handle the problem of byzantine problem [20]. Hyperledger fabric is used in PBFT as a consensus algorithm and it can be handled 2/3 malicious byzantine replica.
A new block is resolute in a round and it is a most important that chosen depends upon the same policy. The policy is responsible for ordering the transactions. In the PBFT, the entire process is separated into three parts such as pre-prepared, prepared and commit.
2.3.2 Proof-of-Work (PoW)

The PoW is a well-known consensus mechanism used in Bitcoin [11]. Each node in the network is computing a hash value of the altering block header. The consensus mechanism wants that the calculated values must smaller or equal to the given value. All participants in the network have to compute a hash value using different nonce value until reach the target. If anyone node achieve relevant value, then all others must confirmed the correctness of the value. The node calculated the hash values is called miner and procedure of PoW is called mining. However, it required a complicated computational process in the authentication process.

2.3.3 Proof-of-Stack (PoS)

PoS is take power consumption and enhance the node performance when compared with PoW [22]. PoS require validating the ownership of the amount of currency instead of trying client to find a nonce in an infinite gap. The PoS protocol randomly provides privilege to create a block between the selected validators based upon the value of their stakes. The selected validator rewarded by the transaction fee partially or entirely. But, it produces inaccuracy during vulnerable attacking. Compared with PoW, PoS has low latency, lower resource usage and computation power.

2.3.4 Delegated PoS (DPoS)

DPoS is a consensus algorithm upholding an unquestionable contract on the fact across the network and transaction validations. Among the validator, the voting power is to determine the peer who fills the role of a delegate for maintaining the network and transaction validations known as forging. The DPoS is to hold stake vote to select by block verifiers or creators.

2.3.5 Ripple

Ripple is a kind of consensus algorithm that employ collectively trusted and sub-networks within the bigger networks [23]. The participating nodes are separated into two types such as server and client. The server performs the contribute process of consensus. The client performs only transferring of funds. In ripple, each server has a unique node list (UNL) that is a significant value to the server to query. The server would query the node in UNL when put a transaction in the ledger. The transaction is crammed into the ledger if the received agreement has reached 80%. The fault node is corrected by the ledger whose the percentage has less than 20%.

2.3.6 Tendermint

Tendermint is a kinds of byzantine consensus algorithm. In which, a new block is resolute in a round manner [24]. A proposer would be chosen to relay an unverified block. Hence, all nodes in the network are required to know the proposed selection. Tendermint could divide into three steps prevote step, pre-commit step and commit step. The validator decide whether to relay a prevote for the planned block in the prevote step. The pre-commit step is performed if the node has obtained more than 2/3 of the planed block, then it relay a pre-commit for that block. If the node has obtained over 2/3 of pre-commits, it goes into the commit steps.
2.4 Characteristic of blockchain technology

The blockchain technology is a decentralized ledger scheme with high security to enhance the integrity and trust of transactions [25]. This section discusses some major characteristics of blockchain as follows,

- **Decentralized**: the decentralized property of the blockchain can manage the single point of failure and scalability. The property of decentralized and distributed ledger in blockchain is reducing the latency.

- **Transparent**: the blockchain technology is produced high level transparency with the help of sharing transaction information between all participants’ users who involving the transactions. The role of a third party is no need to verify the any transaction as in the centralized technology.

- **Immutability**: the blockchain can ensure the every transaction by creating the immutability ledger. In the blockchain, the every transaction in the block is cannot change. Because, the blocks of the every transactions are stored permanently by connecting each other using hash values.

- **Security**: it provides more security for storing the data by using public key infrastructures and protect from changing by the malicious user. The blockchain mechanism protects the entire system by eliminating from the problem single point of failure.

- **Efficiency**: the distributed mechanism is used to enhance the efficiency which verifies the all transaction in the distributed environments. The blockchain is more efficiently compared with a centralized system in terms of completion speed, cost and risk management.

- **Traceability**: the property of timestamps in the blockchain is used to recognize and store each transaction. This helps in participating node to keep the order of transactions for data traceable. The timestamp property can help to reduce the cost of transaction traceability. The traceability of the blockchain system is secure and consistent, and solves the problem of double spending.

- **Anonymity**: The blockchain technology is using the asymmetric encryption techniques for encrypting the data. The encryption of data in the blockchain is used two methodologies including data encryption and digital signatures. Data encryptions are used to make sure the security of transaction data and decrease the possibility of losing transaction data. Transaction data is use the digital signed to indicate the identity of the originality of transaction data.

3. Blockchain based data integrity scheme:

The cloud storage is widely used resource service compared with other resource due to its higher efficiency and low cost. The cloud user can enjoy high-quality storage service for their data. Because the user data store is resides at isolated server. The security issue of CSS is a major role that makes the user worried about their outsourced data. In the earlier stage, the many public data integrity scheme is depending upon a trusted TPA to ensure the integrity of the data.
However, the malicious TPA can collaborate with the CSP to obtain various benefits. In order to mitigate the user tension, the data integrity scheme is very important to verify the integrity of the outsourced data [26]. Hence, the auditability of property of the blockchain is utilized for data integrity checking that the transaction records of CSS can be easily verified and tracked [2]. The blockchain can successfully preserve the data consistency in the distributed computing systems [27]. Also, the blockchain technology is used for verifying the data integrity of stored data in different platforms such as Big data [28], IoT [29], cloud environment [30]. The following paragraphs are discussed some research papers about data integrity inspection of the cloud storage data by using blockchain technology.

Xue et al. (2019) [17] propose an identity based public auditing scheme by using blockchain for maintaining data integrity of outsourced data against various attacks. The authors use the nonce of public blockchain (Bitcoin) for selecting the challenge message of identity-based public auditing (IBPA). Additionally, the auditing results are stored in the public blockchain with traceable and auditable. Security of IBPA is under the oracle model based on Diffie-Hellman problem. The proposed data integrity scheme is produced low communication overhead and low computational cost. Yu et al. (2017) [28] developed a decentralized big data auditing method for the applications of smart city using ability of blockchain to enhanced reliability and strength without the need of a centralized TPA. Wang et al. (2019) [30] developed blockchain-based private PDP data integrity scheme. In which, the data irreversibility property of the blockchain is used to store some metadata for client that can be used to verify the remote data integrity. The proposed method uses the blockchain that produce the corresponding metadata and forwarded it to the client.

Wei et al. (2020) [31] proposes the data integrity protection scheme by using blockchain technology. The developed data integrity scheme is built by virtual machine proxy method and unique hash value corresponding to the file produced by the merkal hash tree (MHT) which is used to examine the changes of files. Zhang et al. (2019) [32] develop certificateless public verification method in opposition to procrastinating auditors (CPVPA) by means of blockchain

### Table 2: Performance comparisons of data integrity schemes

| Ref. No. | Batch | Dynamic | Public | Private | Block less | Certificate less |
|----------|--------|---------|--------|---------|-----------|-----------------|
| [16]     | ✓      |         | ✓      |         |           |                 |
| [17]     | ✓      |         | ✓      |         |           |                 |
| [26]     | ✓      |         | ✓      |         | ✓         |                 |
| [28]     | ✓      | ✓       | ✓      |         | ✓         |                 |
| [30]     |        |         | ✓      |         | ✓         |                 |
| [31]     |        |         | ✓      |         |           |                 |
| [32]     |        |         | ✓      |         |           |                 |
| [33]     | ✓      |         | ✓      |         |           |                 |
| [34]     | ✓      |         | ✓      |         |           |                 |
| [35]     | ✓      |         | ✓      |         |           |                 |
| [36]     |        |         | ✓      |         |           |                 |
| [37]     |        |         | ✓      |         |           |                 |
| [38]     | ✓      |         | ✓      |         |           |                 |
| [39]     | ✓      |         | ✓      |         |           |                 |
| [40]     |        |         | ✓      |         | ✓         |                 |
technology. The key thought is the auditors need to store each verification outcome in a blockchain. Because, the transactions of the blocks are time-sensitive, the verification process can be time-stamped after the related transaction is stored. It can use to verify whether auditors achieve the verifications at the approved time.

Hao et al. (2019) [33] propose a decentralized form based on blockchain that contains the several collaborative verification peers (VPs). The each VPs maintains a replication of the complete blockchain to keep away from malicious tampering. The proposed scheme is allow DO to store and make sure the verification information by writing and get back the blockchain. Also, proposed scheme is extend the algorithm by introducing then verification group (VG) constituting by some VPs controlled by Inter-Group and Inner-Group consensus protocols. Lu et al. (2020) [34] design significant decentralized data integrity auditing technique based on Hyperledger Fabric (HF-Audit). TPA can be dynamically determined for each auditing task by using Hyperledger Fabric as a communication platform. Also, the proposed system design a significant auditing method for data integrity based on bilinear pairing and commitments in order to enhance the scalability.

Wang et al. (2020) [35] proposed a blockchain-based fair payment that adopting blockchain to replace TPA for public cloud storage auditing. In the proposed system, DO and CSP will run a blockchain-based smart contract to ensure the details of transactions. The CSP is necessary to submit data possession proof frequently. The proposed scheme is present no interactive public PDP and design a smart contract using blockchain technology for public cloud storage auditing. Kalis et al. (2018) [36] develop a hash validation technique using blockchain technology. The actual stored data can be authenticated against the blockchain based hash and an employ inside an application check trail to authenticate the audit trail data. This implementation demonstrates that the hash validation is capable to identify malicious and fortuitous changes that were prepared for the data.

Wang et al. (2019) [37] develop a new personal health record sharing techniques with data integrity verification method based on blockchain technology. The proposed method uses the attribute-based encryption (AE) and searchable symmetric encryption (SE) techniques to accomplish privacy protection, fine-grained access control and keyword search.
A blockchain based method is used to supervise propose a bilinear mapping and blockchain based data integrity scheme (BB-DIS) for large-scale IoT data to reduce computational and communication overhead. In the proposed method, large scale IoT data are segmented into shards and homomorphic verifiable tags (HVTs) are produce for sampling confirmation. Then, data integrity method can be accomplished according to the individuality of bilinear mapping and blockchain transactions.

Yue et al. [39] proposes new data integrity checking scheme based blockchain in Peer-to-Peer cloud storage. The present Merkle tree for data integrity checking and investigate the performance of system under dissimilar Merkle trees construction. Also, the authors develop a logical sampling scheme to create sampling checking more efficient. Yuea et al. (2020) [40] propose a new data integrity checking scheme in a decentralized edge-cloud storage (ECS) using blockchain-based framework without TPA. The authors employ the Merkle tree with dynamic demanding records to verify data integrity and optimize the system performance by analysing different Merkle tree structures. Zhao et al. (2020) [16] designed a new remote data

### Table 3: Comparisons of the data integrity scheme

| Ref. No. | Verification Algorithm | Consensus algorithm | Merits | Demerits |
|----------|------------------------|---------------------|--------|----------|
| [16]     | Lifted EC-ElGamal cryptosystem, bilinear pairing | PoW | High accuracy | High computation cost |
| [17]     | Diffie-hellman | Bitcoin system | Low communication overhead and computational cost | Lack of scalability |
| [28]     | Bilinear pairings | PBFT | lower communication and computation costs | Low accuracy |
| [30]     | RSA | Blockchain | More efficient and practical | High computational time |
| [31]     | - | PoW | Produced warning message | Energy consumption is high |
| [32]     | Bilinear maps | PoW | Avoid the certificate management problem with high accuracy | More Energy consumptions |
| [33]     | RSA | VG-WriteBlock, VG-CheckBlock | High security and efficiency | Energy Consumptions |
| [34]     | bilinear pairing and commitments | HF-Audit | High security and scalability | High energy consumptions |
| [35]     | Bilinear Pairings + Diffie-Hellman | PoW | Low execution time | High computational burden |
| [36]     | - | PoC based hash validation function | Can able to distinguish tampering | It cannot prevent tampering |
| [37]     | SE+AE | Ethereum | avoiding the single point failure problem | High computation overhead |
| [38]     | Bilinear Mapping | PBFT | Computational cost and communication overhead | Inability to handle complex data types |
| [39]     | Merkle tree + sampling strategy | Ethereum | Enhancing the performance of data integrity verification scheme. | More power consumptions |
| [38]     | Merkle tree + sampling Strategy with random generation | Ethereum | More flexibility, transparent and auditable | More power consumptions |
integrity verification technique for security of cloud-based information systems. The proposed scheme is combined with leverages the Lifted EC-ElGamal cryptosystem, bilinear mapping and blockchain scheme to carry out proficient public batch signature checking, keep the security and data privacy of the IoT environment. The comparisons of reviewed research paper are shown in Table 2 and 3.

4. Limitations of blockchain based data integrity scheme

DO combine with TPA to achieve the data integrity checking scheme to assess the outsourced data integrity. During the verification process, DO can’t be completely authenticated the consistency of the TPA. Hence, furnish over the checking of data integrity to the malicious and it may direct to data security threats. Table -3 is demonstrates the advantages and disadvantages of existing research papers for problem of data integrity. However, the blockchain based data integrity scheme has been produced higher accuracy when compared with conventional verification scheme. But, the existing research works have to look the troubles of huge computational and communication overhead [38]. The various research works have been developed a blockchain based verification scheme such as [17] , [31] and [35]. In the blockchain, the bitcoin is used PoW consensus protocol to verify the integrity of outsourced data. However, the PoW is taking more energy consumption and consensus process is prohibitively slow for validating [41].

5. Challenges of data integrity verification methods

The data integrity scheme has many challenged during integrity verification using blockchain technology including computation, storage, communication and disk I/O efficiency. The following subsections are discussed about the challenges of data integrity scheme.

• **Computational efficiency:** the pre-processing step is a major concern when data outsourced in the cloud. Proof generation of outsourced data during the data integrity process is taking more computation efficiency. Also, taking more computation efficiency when dealing with meta data.

• **Storage efficiency:** the additional meta data is required along with original data for integrity verification and data recovery. Hence, in this circumstance, the storage efficiency is considered as a very important.

• **Communication efficiency:** Communication effectiveness of the data integrity scheme is considered in three ways such as the original data transfer to the cloud with meta data is overhead, request to the verification process by the data owners, and proof of data integrity is sent reverse to the CSS.

• **Disk I/O efficiency:** Accessing the block for making proof, inconsistent length of block access, amount of blocks, and unordered block indexes are the disturb the efficiency of the data integrity scheme.
6. Conclusions

Data integrity and confidentiality are the major security concerns in a cloud computing environment. Many data integrity models have been developed over the years to concentrate on the data integrity issues. On the other hand, the blockchain technology is a well-known and considerable technique in terms of security issues in various fields. The blockchain technology helps to enhance the performance of data integrity models. The major focus of the paper is to enhance the perception of security issues of data integrity methods based on blockchain technology. The present paper gives a comparative analysis based on characteristics of data integrity schemes and blockchain technology. Also, it recognized the timeline of the development of data integrity schemes based on blockchain technology. Also, it can be identified the future aspects of blockchain-based data integrity scheme. This paper can help to researchers to carry out their future research work. Further, we identify the challenges and security issues of data integrity schemes based on blockchain technology.

7. Future enhancements

In the data integrity scheme, the verification process is significantly essential process. The many research work have been applied a blockchain technology for data integrity verification. In this circumstance, the enhancing performance of the blockchain validation process is challenging and crucial task. Hence, the PoS, DPoS is an alternative consensus algorithm for validating the transaction of blockchain technology. On the other hand, PeerCensus is a well-known consensus algorithm proposed by Decker et al. (2016) [41] that proceed as a certification influence and supervises peer identities in a P2P network, and ultimately improves performance of blockchain. As well, Kraft (2016) proposes a new consensus algorithm that enhance stability of bitcoin performance [42].

References

[1] Nepal S Chen S Yao J and Thilakanathan D 2011 DIaaS: Data integrity as a service in the cloud IEEE 4th International Conference on Cloud Computing, IEEE.
[2] Fan Y Lin X Tan G Zhang Y Dong W and Lei J One secure data integrity verification scheme for cloud storage Future Generation Computer Systems 96 376-385.
[3] Sood S K 2012 A combined approach to ensure data security in cloud computing Journal of Network and Computer Applications 35 1831-1838
[4] Shen W Qin J Yu J Hao R Hu J and Ma J 2019 Data integrity auditing without private key storage for secure cloud storage IEEE Transactions on Cloud Computing.
[5] Zhu Y Hu H Ahn G-J and Yau S S 2012 Efficient audit service outsourcing for data integrity in clouds Journal of Systems and Software 85 1083-1095
[6] Ateniese G Burns R Curtmola R Herring J Kissner L Peterson Z and Song D 2007 Provable data possession at untrusted stores Proceedings of the 14th ACM conference on Computer and communications security
[7] Juels A and Kaliski Jr B S 2007 PORs: Proofs of retrievability for large files Proceedings of the 14th ACM conference on Computer and communications security
[8] Han S Liu S Chen K and Gu D 2014 Proofs of retrievability based on MRD codes International Conference on Information Security Practice and Experience, Springer.
[9] Yuan J and Yu S 2013 Proofs of retrievability with public verifiability and constant communication cost in cloud Proceedings of the 2013 international workshop on Security in cloud computing Editor.

[10] Li A Tan S and Jia Y 2019 A method for achieving provable data integrity in cloud computing The Journal of Supercomputing 75 92-108.

[11] Nakamoto S. 2008 Bitcoin: A peer-to-peer electronic cash system.(2008) Editor.

[12] Casino F Dasaklis T K and Patsakis C 2019 A systematic literature review of blockchain-based applications current status classification and open issues Telematics and Informatics 36 55-81

[13] Yang C Chen X and Xiang Y 2018 Blockchain-based publicly verifiable data deletion scheme for cloud storage Journal of Network and Computer Applications 103 185-193.

[14] Abdullah N Hakansson A and Moradian E 2017 Blockchain based approach to enhance big data authentication in distributed environment 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN) IEEE.

[15] Zhang L Xie Y Zheng Y Xue W Zheng X and Xu X 2020 The challenges and countermeasures of blockchain in finance and economics Systems Research and Behavioral Science 37 691-698.

[16] Zhao Q Chen S Liu Z Baker T and Zhang Y 2020 Blockchain-based privacy-preserving remote data integrity checking scheme for IoT information systems Information Processing & Management 57 102355

[17] Xue J Xu C Zhao J and Ma J 2019 Identity-based public auditing for cloud storage systems against malicious auditors via blockchain Science China Information Sciences 62 32104.

[18] Johnson D Menezes A and Vanstone S 2001 The elliptic curve digital signature algorithm (ECDSA) International journal of information security 1 36-63.

[19] Atlam H F and Wills G B 2018 Technical aspects of blockchain and IoT

[20] Lamport L Shostak R and Pease M 1982 The Byzantine Generals Problem ACM Transactions on Programming Languages and Systems vol 4 no 3 pp. 382-401

[21] Li X Jiang P Chen T Luo X and Wen Q 2020 A survey on the security of blockchain systems Future Generation Computer Systems 107 841-853.

[22] Zamfir V 2015 Introducing casper the friendly ghost Ethereum Blog URL: https://blog.ethereum.org/2015/08/01/introducing-casper-friendly-ghost.

[23] Schwartz D Youngs N and Britto A 2014 The ripple protocol consensus algorithm Ripple Labs Inc White Paper 5

[24] Kwon J 2014 Tendermint: Consensus without mining Draft v. 0.6. fall 1.

[25] Atlam H F and Wills G B 2019 Technical aspects of blockchain and IoT Advances in Computers Elsevier.

[26] Saxena R and Dey S 2019 Data integrity verification: a novel approach for cloud computing Sādhanā 44, 74.

[27] Zhu Z Qi G Zheng M Sun J and Chai Y 2019 Blockchain based consensus checking in decentralized cloud storage Simulation Modelling Practice and Theory 101987.

[28] Yu H Yang Z and Sinnott R O 2018 Decentralized big data auditing for smart city environments leveraging blockchain technology IEEE Access 7 6288-6296
[29] Machado C and Fröhlich A A M 2018 IoT data integrity verification for cyber-physical systems using blockchain,” in 2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC), IEEE.
[30] Wang H Wang Q and He D 2019 Blockchain-Based Private Provable Data Possession IEEE Transactions on Dependable and Secure Computing.
[31] Wei P Wang D Zhao Y Tyagi S K S and Kumar N 2020 Blockchain data-based cloud data integrity protection mechanism Future Generation Computer Systems 102 902-911
[32] Zhang Y Xu C Lin X and Shen X S 2019 Blockchain-based public integrity verification for cloud storage against procrastinating auditors IEEE Transactions on Cloud Computing.
[33] Hao K Xin J Wang Z and Wang G 2020 Outsourced data integrity verification based on blockchain in untrusted environment World Wide Web, 1-24.
[34] Lu N Zhang Y Shi W Kumari S and Choo K-K R 2020 A secure and scalable data integrity auditing scheme based on hyperledger fabric Computers & Security 92 101741
[35] Wang H Qin H Zhao M Wei X Shen H and Susilo W 2020 Blockchain-based fair payment smart contract for public cloud storage auditing Information Sciences 519 348-362.
[36] Kalis R and Belloum A 2018 Validating data integrity with blockchain 2018 IEEE International Conference on Cloud Computing Technology and Science IEEE
[37] Wang S Zhang D and Zhang Y 2019 Blockchain-Based Personal Health Records Sharing Scheme With Data Integrity Verifiable IEEE Access 7 102887-102901.
[38] Wang H and Zhang J 2019 Blockchain Based Data Integrity Verification for Large-Scale IoT Data IEEE Access 7 164996-165006
[39] Yue D Li R Zhang Y Tian W and Peng C 2018 Blockchain based data integrity verification in P2P cloud storage IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), IEEE.
[40] Yue D Li R Zhang Y Tian W and Huang Y 2020 Blockchain-based verification framework for data integrity in edge-cloud storage Journal of Parallel and Distributed Computing
[41] Decker C Seidel J and Wattenhofer R 2016 Bitcoin meets strong consistency Proceedings of the 17th International Conference on Distributed Computing and Networking Editor
[42] Kraft D 2016 Difficulty control for blockchain-based consensus systems Peer-to-Peer Networking and Applications 9 397-413.