A Review on Data Integrity Verification Schemes Based on TPA and Blockchain for Secure Cloud Data Storage

S. Sudersan, V. S. Abhijith, M. Thangavel, and P. Varalakshmi

1 Introduction

Nowadays the demand for cloud-based methodologies that offer storage facilities is increasing rapidly. In addition, as information exchange has increased a lot due to the move to a complete virtual platform in every sector and domain everyone has made their way to store the data in the cloud for better access, collaboration, and reliability. However, the verification of the integrity of data in cloud storage has been challenging. In the real-world scenario the encryption algorithm, the signatures, and other mechanisms and policies ensure data privacy. But when data has moved to a remote server (i.e.) cloud storage, it is being tedious to maintain and manage them. So verification of the data and its integrity in a cloud-based system or storage has very high importance. Blockchain offers unique characteristics of resistance from tamper, a noncentralized approach, maintaining very high consistency and stability, and the ability to trace every transaction. The increase in the peer-to-peer (P2P) working on cloud storage leads to a large amount of exhaustion of the data available. Blockchain technology is decentralized, transparent, and immutable storage and a record that proves it has not been tampered [31, 35, 39]. The mechanism of agreeability or consensus ensures the state of the complete system in the blockchain. Smart Contracts are business logic or the processing logic in the Blockchain-based framework and execute the necessary conditions and instructions given based on the conditions.

S. Sudersan · V. S. Abhijith
Department of Information Technology, Thiagarajar College of Engineering, Madurai, Tamilnadu, India

M. Thangavel (✉)
School of Computing Science and Engineering, Vellore Institute of Technology, Bhopal Campus, Bhopal, India

P. Varalakshmi
Department of Computer Technology, Madras Institute of Technology, Anna University, Chennai, Tamilnadu, India

U. P. Rao et al. (eds.), *Security, Privacy and Data Analytics*, Lecture Notes in Electrical Engineering 848, https://doi.org/10.1007/978-981-16-9089-1_10

Billions of users across the world store data in the cloud. However the integrity of data stored in the cloud is not guaranteed, it by nature assumed that the CSP shall not compromise on integrity, that is not, however, the case always. Several solutions emerged to ensure data integrity, one such solution is to introduce a blockchain. In this survey, we have highlighted various methods for data integrity verifications with and without Blockchain and the issues that plague them and highlighted the problems associated with Third-Party Auditors (TPAs). Several solutions to the data integrity problem have been analyzed, their pros and cons weighed. In our survey, we have also tried to find how the inclusion of blockchain can produce a much more efficient, secure, and viable solution.

2 Issues and Challenges

Many reasons stand as an issue or the cause of compromising the integrity in data storage. The errors include those that are intentional or unintentional. Errors in transfer including modifications can happen during the transfer of data from one device to another. Viruses, malware, and other vulnerabilities of the cyber world could also be an indirect chance of compromising data integrity [26]. Some other issues may include system failures and human errors. In some or for some rare chances may include the compromising of the data in the intention of providing more value to the premium consumers in the expectation of higher profit. This may happen in the case of private clouds that intend to provide better security with a pay increase. Thus with the increase in demand for the independent, remote access of data and services and the incorporation of the CSPs in gaining a high revenue with the help of the available resources and services being offered and shared over to a large community the demand for cloud storage and services has raised to a greater extent [23]. However, all these demands show a greater need and brighter side of the technology, there is still some inherent weakness called trust. A CSP can never be trusted always because of the higher chances of being motivated towards malicious events and activities. As an example, a CSP in the case of storage could analyze the usage of data in the storage and may gradually try to delete unnecessary files or information that the user has not made access to for a certain amount of time. Since there is a large spread of the data being stored and the data is scattered across multiple servers in a widespread geographical area the user is unable to inspect every data link visually especially when the size of the data is very large. Thus as a main area of concern, unauthorized physical access could be one of the major issues being faced specifically in the area of storage.

3 State of the Art

Several solutions have been proposed to solve the problem of data integrity, both involving and excluding blockchain technology.

3.1 Challenges Addressed Without Blockchain

Xiling Luo et al. [28] have proposed an integrity verification scheme or methodology based on Boneh Lynn Shacham type of signature. The proposed system has 3 entities: User, Server, and the TPA and involves 5 functions: Key Generation, Token Generation, Puzzle, Responding, and Proof checking. Operations, Key Generation, Token Generation, Puzzle, Responding, and Proof checking. The public and private keys are generated by the user by running the key generation functionality. When the user wants to check the integrity of his/her data, the user requests the TPA, which in turn creates a challenge and sends it to the CSP. The service provider in turn solves the challenge and returns the proof to the TPA. The TPA on receiving the proof provided by the CSP checks and returns the status to the user.

Shaomin Zhang et al. [44] proposed a remote integrity verification methodology using cloud computing. The proposed scheme involves 3 commodities: Owner of the data (DO), the Provider of the Cloud systems (Cloud Service Provider), Auditor on third-Party policy (Third-Party Auditor). The Provider of the Cloud systems uses the generation of key functionalities to generate secret key pairs. The data proofs are protected by random parameters. A challenge is generated by the TPA and sent to the CSP based on set equations and parameters, to which the CSP responds, which the TPA verifies for an integrity based on the verification equations.

Yuan Zhang et al. [40] have proposed a public data integrity verification mechanism for cloud storage systems. To verify the data integrity, the auditor first generates a challenge message and issues it to the cloud server. With the challenging message, the cloud server generates proof of the information and that is then sent to an auditor. After the information is received, the data integrity is verified by the auditor on checking the proof information. Based on the verification results, the auditor informs the user accordingly.

Yuan Ping et al. [4] have proposed a public data integrity verification scheme for secure cloud storage. The system model comprises Users, Cloud Service Providers (CSPs), and Third-Party Auditors (TPAs). Typical working of the system involves the user sending a verification request to the TPA, which chooses challenge information and sends the data, computes the proof, and sends it back to the third-party auditor, which it further processes by computing the sum of hash values, verifies the proof and then transmits the result to the user.

Annamalai Rajendran et al. [34] have proposed a mechanism for data integrity verification using provable identity-based data possession in cloud storage environments. The proposed integrity verification process involves partitioning the file into

blocks and is stored into data and encoded using AES. The data is encrypted and uploaded to the cloud. The TPA performs the auditing process. The user requests the TPA to verify the integrity of data. The TPA performs a probabilistic check to verify the integrity of the user data. The proposed solution is implemented in a multi-storage cloud so that in case of data corruption, the data blocks can be replaced.

Yunxue Yan et al. [29] have proposed a dynamic data integrity verification scheme based on bloom's filters and lattices. The proposed scheme involves three entities: User, Cloud Service Provider (CSP), and Third-Party Auditor (TPA). Bloom's filter is a highly efficient data structure, which uses a vector to represent a set of members. The main function of Bloom's filter is to determine whether an element belongs to this collection or not. To verify the integrity of his/her data, the user sends an audit request to the TPA, which in turn generates a challenge and sends it to the CSP. The CSP on authentication accepts the challenge and generates the corresponding file signature and returns it to the TPA. The TPA verifies the proof and notifies the user accordingly.

Junfeng et al. [7] have proposed a data integrity verification scheme involving a homomorphic hash function that provides greater privacy, efficiency and handles data dynamically better than other conventional schemes. The experimental verification and the security analysis also prove that the system is completely efficient.

Filipe Apolinário et al. [30] have developed S AUDIT which is a service that performs data verification of information stored in cloud servers. The proposed system performs homomorphic verification using digital signatures to prevent retrieval of the data that is protected, especially in the case of cloud-based storage systems. Experimental evaluation reveals that the S-AUDIT-based approach is 7.1% cheaper than using other signature methods like RSA.

Shivaraj Hiremath et al. [36] provides an efficient public audit methodology using TPA to assure integrity. The analysis of the work proves that the approach is safe and the TPA takes only unit time to help to the audit of data.

3.2 Challenges Addressed with Blockchain

Muhammad Saqib Niaz et al. [43] have proposed Merkle hash tree-based techniques to ensure the integrity of outsourced data. The authors consider that there exists a methodology where the system takes the efficiency to securely transfer the data and information with the Data Owner (DO) and the clients. The transmitted information could be the public key of the DO or some hash data. The authors have used deterministic approaches based on Authenticated Data Structures (ADS) to verify the integrity of data (Merkle Hash Tree-based Integrity techniques). Merkle Hash Tree In the signature scheme each end node called leaf holds the current data block hash. Internal nodes hold the hash of the concatenated hashes of their children. In the B + tree, the node that is at the root is either considered as leaf or internal. These nodes hold only the key and not the exact information. Data stays always in the leaf nodes. Leaf nodes form some kind of linked list, which allows for the sequential

traversal of the data. Data integrity schemes based on MHT have been designed by replacing B + trees in Merkle's original signature scheme. The authors have also explained MHT storage in the database and Authenticated Data Extraction (ADE) and have also performed a detailed analysis of the methodologies proposed.

PengCheng Wei et al. [8] have proposed a cloud-based integrity verification and protection mechanism using blockchain. The proposed mechanism involves a virtual agent using cloud and mobile technology. The virtual machine agent mechanism ensures reliable data storage, verification, and monitoring. The integrity protection framework involves a proxy virtual machine model and a unique hash value of the file generated by the Merkle tree, which listens for any changes on the data using smart contracts. In case of any data tampering, a warning is issued to the user.

Pei Huang et al. [9] have proposed a Collaborative Auditing Blockchain (CAB) to ensure data integrity in cloud storage systems. The proposed framework involves consists of consensus nodes, which substitute for the single TPA, in executing audit delegations and recording them. The authors have designed a new data structure ACT (Auxiliary Chain Table, which is employed by each Data Owner (DO) to provide for fast and secure data retrieval and verification. The ACT is a two-dimensional data structure. The proposed system consists of four entities: Data Owners (DO), Private Key generators (PKG), Group Managers (GMs), and Cloud Service Providers (CSPs). The PKG is responsible for setting parameters for the system and generating keys for the GM. The DO has limited computational capabilities and is responsible for generating and sending auditing challenges on-demand to the CSP and tracks changes in the ACT. The GM is designated collectively by a group of Dos, however, GMs, in general, possess more computational powers than Dos. The verification process is triggered by the DO, when it initiates an integrity request to the CSP through the GM, once the CSP authenticates and responds to the request, the CAB performs the consensus process and records the result as a block in the blockchain.

Ahsan Manzoor et al. [3] have proposed an IoT data sharing platform based on blockchain, which uses a proxy re-encryption scheme for securely transmitting the information. The proposed system involves a proxy re-encryption scheme for securely transmitting the information. The collected data is encrypted and stored in the cloud. The data is shared between sensors and data users without the involvement of TPA, using runtime smart contracts between them. The proposed scheme is efficient, secure, and fast and serves as a safe platform for trading, storing, and managing sensor data. The system has been deployed using commercially available sensors and IoT devices.

Xiaodong Yang et al. [10] have proposed a multi-cloud and multi replica auditing scheme based on blockchain. The proposed scheme involves a hash table and modification record table, which dynamically updates the results of group audits. The inherent unpredictability of blockchain is used to construct challenges, which prevents malicious TPA and cloud servers from colluding. Every audit result is recorded in the blockchain making the whole process transparent and efficient.

Dongdong Yue et al. [20] have proposed a data integrity verification methodology based on integrity in P2P remote storage based on BC (blockchain). The proposed framework involves three entities Clients, Cloud Storage Servers (CSS), Blockchain

(BC). The workflow consists of two stages viz. Preparation and verification. The first stage has 5 steps. The verification phase also includes five steps, At first, a challenge is generated by the client and sent to the CSS, which then chooses a shard to verify. Then the CSS calculates the hash digest using a hash function based on the challenge and the selected shard. Then the CSS sends the digest and the auxiliary information to the blockchain. Then using smart contracts the blockchain will compute a new hash root and compare it with the previous root. If they are equal, then the integrity of the data is validated, else the data has been compromised.

Gaetani et al. [5] proposed the methodology of a database that could help in the verification of integrity in the cloud with the help of BC. In this paper, the authors have focused on taking up a real-world concern of a SUNFISH project thereby securing the data in the cloud-based storage systems. They have detailed the research and innovation gaps and the tediousness in explaining and documenting them. They have outlined the design using an effective BC approach.

Kun Hao et al. [6] have proposed a data integrity verification mechanism based on blockchain. The proposed solution involves a blockchain model in a non-trustable environment, called Decentralized Collaborative Verification, which proposes an efficient algorithm called Decentralized Integrity Verification which includes two parts, viz., Write Block, and the Check Block.

Igor Zikratov et al. [37] detailed a blockchain-based methodology of verification of data and integrity. The proposed model involves two components viz. The user side and the server-side. The user side consists of the UI etc. The server-side handles session management, authentication, Handling of transactions, and integrity check. Information and its integrity are verified using cross-checking hashes with blockchain.

Hao Wang et al. [11] have proposed a blockchain-based fair payment auditing for public cloud storage. The proposed Non-Interactive-Provable Data Possession (NI-PPDP) scheme consists of three entities: CSP, data owner, and integrity verifier. The scheme consists of four algorithms and is divided into two phases viz. the setup and auditing phase. The setup phase consists of the key and tag generation phase. The audit phase consists of the proof generation and verification phase.

Chao Wang et al. [24] have proposed a blockchain-based audit and access control mechanism in service collaboration. The proposed solution involves a blockchain for recording the file activities of the users in sequential order. The audit process involves a sequence of steps: The first phase is the startup phase, wherein the user generates a key and broadcasts it to the CSP and auditors. In the second phase, the user computes a verifiable homomorphic digest for every data block in a file and each file is construed as an aggregation of the various homomorphic digests. Then the user stores the aggregation and the file in the audit node and cloud storage respectively. In the verification phase, proof of ownership for the data blocks in the file is requested by the audit node. Then a challenge is generated by the audit node and sent to the Cloud Storage Service Provider (CSSP). In the fourth step, the CSSP builds the data possession certificate and sends it to the audit node. Finally, the audit node verifies the result of the challenge and sends the result to the user.

Xuanmei Qin et al. [12] have proposed a blockchain-based access control scheme for secure cloud data sharing with multiple attributes. The proposed system model consists of five entities viz: Certificate Authorities (CA), Attribute Authorities (Aas), Cloud Service Providers (CSPs), Data Owner (DO), Data User (DU). The entities participating in a blockchain network, which records all transactions occurring between the different entities. The blockchain also assists in performing partially trusted computing and in effectively managing user attributes. The proposed mechanism consists of four phases: System initialization, Encryption phase, token generation phase, and decryption phase. The proposed scheme implements cross-domain management of attributes by using smart contracts and also fosters mutual trust among multiple Aas.

Huaqun Wang et al. [41] have proposed a remote data integrity verification mechanism in a cloud-dependent Health Internet of Things. The proposed Integrity Checking & Sharing (ICS) system consists of four entities: Public Cloud Server (PCS), Hospital, Patient, and Patient's authorized entry set (AuthSet). The proposed is an ensemble of seven linear time algorithms.

Lei Zhou et al. [33] have surveyed data integrity verification of the outsourced big data in the cloud environment. In this paper, the authors have reviewed the state of art DIV (data Integrity Verification) efforts to ensure the integrity of data stored in CSS (Cloud Storage Server). Overhead on the users' side. A clear and detailed classification of the forthcoming DIV approaches based on the user mode and storage type has been presented. Some open problems and challenges have been discussed and several valuable ideas have been suggested for further investigation.

Benil et al. [13] have proposed a cloud-based security scheme in E-health systems using blockchain. The proposed Elliptical Curve Certificateless Aggregate Cryptography Signature scheme (EC-ACS) scheme to protect the integrity of Electronic Health Records (EHR) stored in the Medical Cloud Server (MCS) using public auditing and verification. The patient data is encrypted using Elliptic Curve Cryptography (ECC) and the digital signature is generated using the Certificateless Aggregate Signature Scheme (CAS). The proposed scheme offers integrity, traceability, security, and efficiency.

Haiping et al. [14] developed a blockchain-based scheme for preserving the privacy and security in medical data being shared. The privacy concern towards the healthcare data is being considered now as one of the most sensitive sources of information available. The authors have clearly stated the need and concern for Zero-knowledge proof, proxy re-encryption, bilinear maps in such a model proposed. The proposed model behavior has been categorized as if or not the patient believes in the data meeting the requirements provided by the smart contracts and the perspective of the authority and commercial interests concerning the transactions between the patients and the research institutions. A series of 3 major properties have been detailed concerning providing the utmost privacy protection to the data being transacted. A step-wise mathematical proof and steps of the proposed approach give a great insight into the proposed scheme. A brief analysis of the proposed approach has been given in various contexts such as confidentiality, privacy-preserving, integrity, traceability, and a detailed evaluation regarding the procedure.

Bobo et al. [16] have proposed a model for communication that is claimed to be efficient and reliable in multi-tenant edge clouds with a blockchain-enhanced paradigm. A detailed insight into broker-based enabling. The authors provide a detailed analysis and find the ambiguities like issues in authorization and handling sensitive information like metadata. The proposed solution led to the implementation of Kafka and EOS-based solutions using BC (blockchain). An intensive analysis of the proposed solution also proves the suitability of practice by BPS.

Lei et al. [17] opened up the extension of blockchain-based platforms that help in data integrity in the domain of agriculture. The authors have specifically focussed on a platform for fish farms. The main motive of the paper has been to provide fish farmers a secure storage platform for storing the agriculture data with the help of smart contracts. The proposed methodology is implemented using the proof of concept with the help of Hyperledger fabric in the blockchain. A series of experiments were carried out for the analysis of the proposed methodology which has been demonstrated concerning the efficiency and usability of the system.

Bin et al. [38] has introduced the challenges in ensuring data integrity for IoT-based data in the cloud and have given the reason as the dynamic and inherent nature of the data. The authors also claim the inefficiency of Third-Party Auditor (TPA) based data integrity frameworks with reliability being a factor of concern. Hence the authors implement relevant protocols and a prototype system and the necessary analysis of the system has been carried out. The authors identified protocols like DOA to CSS-Y, DOA to CSS-N, DCAs to CSS-Y, DCAs to CSS-N, where DOA → Owner of the application, CSP → Providers who provision Cloud-based storage, CSS → Cloud Storage Service, DCA → Data Consumer Apps, P2PPS → p2p System, DIS → service for integrating the data, DISSC → Integrity of the data with smart control.

Haiyan Wang et al. [25] aimed at providing data integrity verification in the case of large-scale IoT data. The proposed work implements a BC (blockchain) and Bilinear scheme for Data Integrity. The performance analysis including feasibility, security, dynamicity, and complexity of the BB-DIS system to implement the verification scheme supports the achievement of data integrity. The smart contracts verification process is carried out using the SC-VERIFICATION ALGORITHM and the performance analysis has been compared with other similar models and schemes and a detailed simulation has been presented.

Huang et al. [15] proposed a blockchain-based E-Health system called BCES that helps in handling the issues of handling the manipulation of electronic health records (EHR) and that can be audited. Each query made in the EHR will be written in the Blockchain for traceability and helps in permanent storage. Time-based attributes have also been proposed in addition to the main methodology that provides a proxy-based re-encryption to achieve the grained access control of the most sensitive information. The reason for the implementation in the blockchain is the tamper resistance and the characteristics of providing traceability. The security analysis and performance evaluation of the system has been demonstrated which proves the system to be secure and efficient.

Quanyu et al. [18] utilize blockchain technology to implement a scheme that helps in providing remote data integrity for IoT. The main difference of this approach

from other existing approaches is that the scheme does not involve any other third parties. The scheme proposes the usage of Lifted Elliptic Curve El-Gamal cryptography, bilinear pairing with the help of BC (blockchain) to provide batch signature verification thus protecting and securing the devices in an IoT-based system.

Jiaxing et al. [19] utilizes the blockchain-based approach to develop an auditing scheme that helps in verifying the data integrity in the case of cloud storage. The scheme is claimed to be different from other existing approaches as it consists of three participatory entities out of which only two entities are said to be predefined entities. The two predefined entities could be the data owner (DO) and the cloud service provider (CSP) and the main assumption of the system is that they don't have the possibility of trusting the other. Also, there isn't any role for the TPA. A hashtag-based methodology is also being followed. A series of experimental results have been demonstrated that could provide computation and communication.

Yuan et al. [27] proposed a verification methodology that does not require a certificate wherein procrastinating auditors using blockchain. The proposed scheme or methodology involves the auditors recording each verification result into the blockchain in the form of a transaction. The verifications are also time stamped when they are recorded in the blockchain. The paper also demonstrates rigorous security proofs and a comprehensive valuation of the performance proving the system to be efficient.

Dongdong et al. [32] developed a blockchain-based framework without the help of Third-Party Auditors for data integrity verification. The scenario has been designed to take place in a decentralized edge cloud storage environment. The Merkle tree with random challenging numbers has been employed. Hence to solve the problem of these like the limitation in the number of resources available can be solved by using certain rational sampling verification strategies as explained by the authors.

Velmurugadass et al. [21] creates a SDN. It consists of mobile-based IoT devices, open flow, and controllers based on blockchain, servers, servers for authentication (AS), an investigation. It uses an elliptic curve algorithm and is transferred to the cloud. The control system helps in maintaining the evidence and signatures based on the SHA 256 algorithm. The experimental analysis reveals that it gained better performance in response, accuracy, throughput, and security arguments.

Ren et al. [1] identifies and proposes a scheme based on identity proxy using a signature that is aggregated to enhance efficiency and reduce the space usage and bandwidth. The methodology also proves to be cost-efficient in the case of other ordinary signature schemes and the performance is also very high.

Abdullah et al. [2] present a health care management of data that helps in increasing the privacy of data using blockchain methodologies. Cryptosystems and operations are being mentioned in the encryption of data and which helps in ensuring pseudonymity. The analysis of the proposed schemes along with the data processing procedures claims to be cost-effective.

Rupa et al. [22] have proposed a blockchain-based solution to ensure the integrity and security of VC-based device data. The proposed system is implemented by deploying an IoT-based application in a vehicle monitoring system. The various device data such as technical information, vehicle reactions are stored in the cloud

storage. Pentatope based Elliptic Curve Cryptography and SHA is used to ensure privacy. On security and performance analysis the proposed scheme is proved to be efficient and secure.

Shangping et al. [42] have proposed a new blockchain-based data integrity verification scheme for ensuring integrity in health records. The usage of blockchain in the proposed scheme averts the single point of failure problem and ensures fault tolerance. Compared with other existing schemes, the proposed scheme allows a safer environment for sharing the private attributes, keys, etc. of users. The scheme uses smart contracts to evaluate the integrity of the stored data. The security and performance analysis of the proposed solution proves that the scheme is secure, efficient, and feasible for commercial use.

4 Inferences of the Review

Storage, computing, and services are the main resources of cloud computing. In the advanced development of cloud computing technologies, security has been a major concern and more specifically data integrity verification. Several methodologies have been proposed with and without Third-Party Auditors (TPA). It cannot be guaranteed that the TPA would legitimately cooperate with Cloud Service Providers (CSPs) or may be subject to flattery due to some interest or for some other reasons. Moreover, there are chances of loss of confidentiality on the part of CSPs due to some special interest towards premium customers in the private cloud scenario. Even with the development and advanced computing techniques, the security and confidentiality of encryption are under serious threat. Hence, there is a necessity that feasibility should go hand in hand with technology in developing integrity. Verification service that avoids the problem of untrustworthy TPAs and CSPs. However, Blockchain has been identified to be one of the most trustworthy decentralized and transparent systems which have been thought of as a solution to the problem and to overcome the constraints of TPA thus providing authentication, data recovery, backup, and confidentiality in addition to data integrity verification. The main aim and objective are to understand the issues, specifically the proposals made in the area to solve the issues both involving blockchain and to contrast those without involving blockchain. The author aims to highlight the pros and cons involved in the use of blockchain in solving the problem. As an inference of the survey, it has been found emphasized that involving Blockchain technology in solving the problem is the best approach in solving the problem because of the series of validation networks that the blockchain has and the additional feature of data traceability and recoverability. The main problem addressed by many researchers is the loss of confidentiality when TPAs are involved even if the data has been encrypted. The validation of every transaction in the blockchain with its characteristic of being immutable that each transaction or data entered in the blockchain can never be modified or changed to an extent gives a partial solution to the problem thereby replacing the need of the TPAs. Besides, there are a lot of technologies that Blockchain follows to develop its application

like the Kafka in Hyperledger fabric which is believed to bring an entire change in the existing models and constraints. With all these advancements and cons being negligible when compared the blockchain-based implementation of the data integrity service schemes proves to be a lot more trustworthy and secure since there is less involvement of the validators being able to see the transactions that they validate and traceability that if any modifications are done would be explicitly be caught. And this would develop a fear against all those illegitimate activities from the side of CSPs thus providing a safer solution to the issue concerned.

5 Conclusion and Future Scope

Even after all the advancements in security and technologies there still stands a concern of data integrity. As and when technology evolves and new methodologies are being proposed, the changes in the system being vulnerable are high. Still, loopholes are being generated every time and new methodologies are being proposed to handle them. In this paper, we have undertaken an extensive survey on the various proposals, improvements, and suggestions that could help in verifying the data integrity in cloud-based storage systems involving both blockchains versus non-blockchain-based solutions, analyzing their respective complexities, efficiency, feasibility, etc. As we have come to know through this survey, blockchain-based solutions are more efficient and secure. We intend to extend our work to building a secure system. The focus is also on the comparison of secureness in using Blockchain as a technology for ensuring data integrity-based issues, especially in cloud-based environments.

References

1. Ren Y, Leng Y, Qi J, Sharma PK, Wang J, Almakhadmeh Z, Tolba A (2021) Multiple cloud storage mechanisms based on blockchain in smart homes. Futur Gener Comput Syst 115:304–313
2. Omar A, Md, Bhuiyan, Basu A, Kiyomoto S, Rahman S (2019) Privacy-friendly platform for healthcare data in cloud-based on blockchain environment. Futur Gener Comput Syst 95C:511–521
3. Manzoor A, Braeken A, Kanhere SS, Ylianttila M, Liyanage M (2021) Proxy re-encryption enabled secure and anonymous IoT data sharing platform based on blockchain. J Netw Comput Appl 176
4. Ping Y, Zhan Y, Lu K, Wang B (2020) Public data integrity verification scheme for secure cloud storage. Information 11(9)
5. Gaetani E, Aniello L, Baldoni R, Lombardi F, Margheri A, Sassone V (2017) Blockchain-based database to ensure data integrity in cloud computing environments. ITASEC
6. Hao K, Xin J, Wang Z, Wang G (2020) Outsourced data integrity verification based on blockchain in untrusted environment. World Wide Web 23
7. Tian J, Jing X (2020) Cloud data integrity verification scheme for associated tags. Comput Secur 95

8. Wei P, Wang D, Zhao SK, Tyagi S, Kumar N (2020) Blockchain data-based cloud data integrity protection mechanism. Futur Gener Comput Syst 102:902–911
9. Huang P, Fan K, Yang H, Zhang K, Li H, Yang Y (2020) A collaborative auditing blockchain for trustworthy data integrity in cloud storage system. IEEE Access
10. Yang X, Pei X, Wang M, Li T, Wang C (2020) Multi-replica and multi-cloud data public audit scheme based on blockchain. IEEE Access
11. Wanga H, Qin H, Zhao M, Wei X, Shen H, Susilo W (2020) Blockchain-based fair payment smart contract for public cloud storage auditing. Elsevier, Information Sciences
12. Qin X, Huang Y, Yang Z, Li X (2020) A blockchain-based access control scheme with multiple attribute authorities for secure cloud data sharing. J Syst Arch 112
13. Benil T, Jasper J (2020) Cloud-based security on outsourcing using blockchain in E-health systems. Comput Netw 178
14. Huang H, Zhu P, Xiao F, Sun X, Huang Q (2020) A blockchain-based scheme for privacy-preserving and secure sharing of medical data. Comput Secur 99
15. Huang H, Sun X, Xiao F, Zhu P, Wang W (2021) Blockchain-based eHealth system for auditable EHRs manipulation in cloud environments. J Parallel Distrib Comput 148:46–57
16. Huang B, Zhang R, Lu Z, Zhang Y, Wu J, Zhan L, Hung PC (2020) BPS: a reliable and efficient pub/sub communication model with blockchain-enhanced paradigm in multi-tenant edge cloud. J Parallel Distrib Comput 143
17. Hang L, Ullah I, Kim DH (2020) A secure fish farm platform based on blockchain for agriculture data integrity. Comput Electron Agric 170
18. Zhao Q, Chen S, Liu Z, Baker T, Zhang Y (2020) Blockchain-based privacy-preserving remote data integrity checking scheme for IoT information systems. Inf Process Manag 57
19. Li J, Wu J, Jiang G, Srikanthan T (2020) Blockchain-based public auditing for big data in cloud storage. Inf Process Manag 7(6)
20. Yue D, Li R, Zhang Y, Tian W, Huang Y (2020) Blockchain-based verification framework for data integrity in edge-cloud storage. J Parallel Distrib Comput 146:1–14
21. Velmurugadass P, Dhanasekaran S, Shasi Anand S, Vasudevan V (2020) Enhancing blockchain security in cloud computing with IoT environment using ECIS and cryptography hash algorithm. Mater Today: Proc 37:2653–2659
22. Ch R, Srivastava G, Gadekallu TR, Maddikunta PK, Bhattacharya S (2020) Security and privacy of UAV data using blockchain technology. J Inf Secur Appl 55
23. Wang Y, Chen Z, Wang K, Yang Z (2019) Education cloud data integrity verification based on mapping-trie tree. In: International conference on machine learning, big data and business intelligence, vol 1, pp 155-158
24. Wang C, Chen S, Feng Z, Jiang Y, Xue X (2019) Block chain-based data audit and access control mechanism in service collaboration. In: IEEE international conference on web services, pp 214–218
25. Wang H, Zhang J (2019) Blockchain-based data integrity verification for large-scale IoT data. IEEE Access
26. Sharma P, Jindal R, Borah MD (2019)Blockchain-based integrity protection system for cloud storage. In: 4th technology innovation management and engineering science international conference (TIMES-iCON), pp 1–5
27. Zhang Y, Xu C, Lin X, Shen X (2019) Blockchain-based public integrity verification for cloud storage against procrastinating auditors. IEEE Trans Cloud Comput 9(3):923–937
28. Luo X, Zhou Z, Zhong L, Mao J, Chen C (2018) An effective integrity verification scheme of cloud data based on bls signature. Secur Commun Netw
29. Yan Y, Wu L, Gao G, Wang H, Xu W (2018) A dynamic integrity verification scheme of cloud storage data based on lattice and Bloom filter. J Inf Secur Appl 39:10–18
30. Apolinário F, Pardal M, Correia M (2018) S-audit: efficient data integrity verification for cloud storage. In: 17th IEEE international conference on trust, security and privacy in computing and communications/12th IEEE international conference on big data science and engineering, pp 465–474

31. Krithikashree L, Manisha S, Sujithra M (2018) Audit cloud: ensuring data integrity for mobile devices in cloud storage. In: 9th international conference on computing, communication and networking technologies (ICCCNT), pp 1–5
32. Yue D, Li R, Zhang Y, Tian W, Peng C (2018)Blockchain-based data integrity verification in p2p cloud storage. In: Proceedings of the IEEE 24th international conference on parallel and distributed systems (ICPADS'18), pp 561–568
33. Zhou L, Fu A, Yu S, Su M, Kuang B (2018) Data integrity verification of the outsourced big data in the cloud environment: a survey. J Netw Comput Appl
34. Rajendran A, Balasubramanian V, Mala T (2017) Integrity verification using Identity-based provable data possession in multi-storage cloud. In: International conference on computational intelligence in data science (ICCIDS), pp 1–4
35. Ferretti L, Marchetti M, Andreolini M, Colajanni M (2017) A symmetric cryptographic scheme for data integrity verification in cloud databases. Inf Sci 422
36. Hiremath S, Kunte S (2017) A novel data auditing approach to achieve data privacy and data integrity in cloud computing. In: International conference on electrical, electronics, communication, computer, and optimization techniques (ICEECCOT), pp 306–310
37. Zikratov I, Kuzmin A, Akimenko V, Niculichev V, Yalansky L (2017) Ensuring data integrity using blockchain technology. In: Proceedings of the 20th conference of open innovations association, pp 534–539
38. Liu B, Yu XL, Chen S, Xu X, Zhu L (2017) Blockchain-based data integrity service framework for IoT data. In: IEEE international conference on web services (ICWS)
39. Lin C, Shen Z, Chen Q, Sheldon FT (2016) A data integrity verification scheme in mobile cloud computing. J Netw Comput Appl 77
40. Zhang Y, Xu C, Li H, Liang X (2016) Cryptographic public verification of data integrity for cloud storage systems. IEEE Cloud Comput
41. Wang H, Li K, Ota K, Shen J (2016) Remote data integrity checking and sharing in cloud-based health internet of things. IEICE Trans Inf Syst E99.D:1966–1973
42. Wang S, Zhang D, Zhang Y (2016) Blockchain-based personal health records sharing scheme with data integrity verifiable. IEEE Access
43. Niaz MS, Saake G (2015) Merkle hash tree-based techniques for data integrity of outsourced data. In: 27th GI-workshop on foundations of databases
44. Zhang SM, Xu YC, Wang BY, Xiao J, Niu R (2014) A remote data integrity verification scheme based on cloud computing. Appl Mech Mater 9:644-650