

Block Chain - Assisted Secure Document Sharing System

Lokesh S

Department of Information Security
School of Computer Science and Engineering
Vellore Institute of Technology,
Vellore-632014, Tamilnadu, India

Dr. Manikandan K

Professor Grade 1
School of Computer Science and Engineering
Vellore Institute of Technology,
Vellore-632014, Tamilnadu, India

Abstract - Secure document sharing and verification is required in areas where data integrity and authenticity is required such as the law, healthcare, and governmental services. Traditional centralized systems are not visible on traceability and can be easily altered and tampered with by unlawful hackers.

The article introduces a Blockchain-Assisted Secure Document Sharing System on the basis of a multi-chain blockchain architecture and cryptographic techniques to ensure the integrity, validity, and verifiable trace of documents. The design assumes the use of both RSA-based digital signatures to verify the sender and prevent repudiation and SHA-256 hash to give a unique fingerprint to each document, the transactions of the sidechain are stored in an architecture referred to as the sidechain and to ensure immutability and effective scaling, the sidechain hash is anchored by a main chain periodically. Although the main chain maintains an unchanging reference of sidechain states to enhance security and auditability, each block in the sidechain stores the document hash, timestamp as well as the last block reference to form a tamper-resistant registry. Also, the proposed system supports the chain- f-custody, through which the authorized institutions will verify the authenticity of documents without revealing its original content. The system is more secure, transparent, and more integrity-based compared to conventional document management systems, and hence the experimental implementation demonstrates that the system can be utilized in sharing sensitive documents in institutions and courts.

I. INTRODUCTION

Secure document sharing is essential in most areas, such as courts, government establishment, medical services, and companies. One of the challenges is the

assurance of the validity, integrity, and traceability of the numerous sensitive papers that are distributed in digital form. The traditional document management systems are based on the central databases that are prone to failure of the system, manipulation of the data, and unauthorized access. These issues render the digital document sharing systems less credible. The blockchain technology is a reliable solution since it is decentralized and immutable. A blockchain provides integrity and transparency of stored data because it is extremely hard to alter it once it has been stored. Due to this, it is possible to save safe records of transactions of documents using blockchain. Cryptography is also useful in ensuring that the authenticity of documents is protected. Whereas RSA digital signatures ensure identification and non-repudiation of the sender, the results of a hash in SHA-256 are unique to a particular document and hence integrity is to be checked. This paper introduces a Blockchain-based Secure Document Sharing System, which uses the digital signatures, multi-chain, and cryptographic hashing. Although the sidechain hashes are attached by the main chain to provide additional security and impenetrability, a sidechain records document hash transaction. The given solution enhances transparency, security, and trust in sharing digital documents.

II. RELATED WORK

Over the last several years, several blockchain-based solutions based on the secure management of documents and their verification have been explored. To ensure integrity of the documents and detect unauthorized modifications, primitive schemes focused on an attempt to use cryptographic hashing functions. Some systems used SHA-based hashing methods to generate unique

prints of documents and verify their authenticity when transmission was being done [1]. To effect authentication and non-repudiation in the document sharing scenarios, development subsequently came up with digital signature algorithms such as RSA and other public-key cryptography methods [2]. The emergence of blockchain technology prompted researchers to integrate tamper-resistant documentation verification and safe record keeping based on the use of distributed ledger systems. Document management systems with blockchain have been proposed to increase the level of transparency in sharing data and maintaining immutable records of document transactions [3].

To enhance scalability, efficacy, and maintain a strong security assurance, more current approaches have explored the use of multi-chain or sidechain designs [4]. These systems allow document hashes to be stored in sidechains by anchoring their verification records to a primary blockchain. Most of the existing systems are yet to get scalability, efficient verification, and secure chain-of-custody tracking in place even though these solutions improve data integrity and traceability. To achieve the desired results of providing safe, verifiable, and tamper-resistant document exchange, the proposed solution combines RSA digital signatures, SHA-256 hash, and a multi-chain blockchain design.

III. METHODOLOGY

A. Architecture Diagram

The three major components of the proposed architecture are the Doctor System, the Blockchain Network, and the Court Verification System. The user signs in and scans his face and blink detection as liveness verification to the doctor system before recording video, audio, and report data as digital proof. The data obtained is then hashed using SHA-256 to generate a hash value that is a digital fingerprint of the evidence. An RSA digital signature is then employed to authenticate the hash to verify the validity of the source. The generated hash values are stored in a blockchain network with a multi-chain design which consists of a sidechain and a main blockchain. The sidechain records are anchored to the main blockchain to guarantee security and immutability, whereas the hashes of the documents and media are stored in the sidechain.

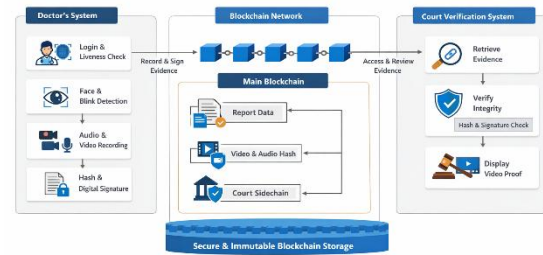


Fig 1. Architecture Diagram

The stored documents within the Court Verification System are accessed by the authorized officials and verified with the help of digital signatures and the hash values to verify the evidence. The validated video or document evidence can be presented to be examined by a legal expert in case the hashes are the same, and the evidence proves to be valid and intact.

The blockchain structure also maintains a chain of custody in which every transaction has been linked with hashes of the previous block and is time stamped. This provides a reliable way of administering digital evidence safely in institutional and legal contexts by ensuring that all the records of evidence remain readable, traceable, and inaccessible.

B. Proposed Algorithm

The proposed Secure Document Sharing Algorithm based on Blockchain supports, provides secure and efficient, verifiable, and tamper-resistant document management with digital signatures of RSA, hash, and multi-chain blockchain architecture, as well as a combination of these tools. In order to prevent identity faking, the user has to be first authenticated through the use of login and liveness checking. Audio, video, and report documentations are captured after a user has been authenticated. Each file is processed by the SHA-256 hashing algorithm in order to obtain a unique hash value which acts as a digital fingerprint of the data. The hash thus formed is signed using RSA private key to form a digital signature that grants authenticity and non-repudiation. To ensure the secured ledger, the signed hash is then stored on the sidechain, where the blocks are linked through the previous block hash.

In a bid to make the stored records even more immutable and avoid manipulation of the stored records, the final hash of the sidechain is periodically appended on to the main blockchain. The system will get the digital signature and saved hash by the blockchain during verification. The document is rehashed using SHA-256 and compared with the value of the hash that had been stored. The document integrity and validity is checked when the hashes are equal and when the RSA signature is authentic. This process allows the secure verification by the authorized agencies and ensures the reliable chain of custody of digital evidence.

M → Input Document / Evidence
 H → SHA-256 Hash Value
 K_{priv} → RSA Private Key
 K_{pub} → RSA Public Key
 S → Digital Signature
 SC → Sidechain
 MC → Main Chain

Input: digital evidence (audio, video, report document).

1. User Authentication: Once the user has logged in, he or she is able to use face and blink detection which requires him/her to prove that he/she is alive.
2. Gathering Evidence: Obtain electronic evidence, audio, video, various report data.
3. Generation of hash: Calculating a unique hash value: Using hash Sha256: $H = SHA-256(M)$
4. Creation of Digital Signatures: Say the resulting hash using the RSA private key: $S = RSA_sign(H, K_{priv})$
5. Record the metadata in a separate block which is known as a sidechain block and store the document hash in it. $Block = \{Index, Timestamp, Data Hash, Previous Hash\}$
6. Sidechain Linking: In order to make the chain unalterable, one is to bind every new block with the hash of the previous block.
7. Main Chain Anchoring: In order to ensure permanence and global verification, periodically anchor the latest sidechain hash to the main blockchain.

8. Document Verification: Get the stored hash of the document in the blockchain and reproduce it. It is proved to be an integrity when $H_{computed} = H_{stored}$
9. Signature Checking: In order to verify the digital signature, use the RSA public key. $Verify(S, H, K_{pub})$
10. Logging and Completion: Record documents of chain of custody and store findings of verification.
- 11.

Output: Authenticated and verifiable blockchain record and integrity of document.

C. Working Algorithm

Upon the entry of an authorized user into the system and the recording of digital evidence, such as audio, video, and report papers, the process begins. The system is configured to verify a user identity and verify liveness before processing the data to ensure that the input is of a real user. Successful verification is then converted into a standardized digital form of the captured evidence to be subject to processing by the cryptographic modules. The accepted evidence is represented as an input M.

$$M = \{video, audio, report\}$$

Hash Generation Process

It performs the hash computation with the help of the SHA-256 hashing algorithm to calculate a different hash value after receiving the input document M. This hash is the digital signature of the document and ensures that so simple alterations in the data will lead to a completely different hash value. The generation of the hashes at the stage of generating the hash is the guarantee of document integrity, without storing the original data in the blockchain.

$$H = SHA-256(M)$$

where

M → Input document or evidence

H → Generated hash value.

Digital Signature Process

The generated hash value is authenticated and signed with the RSA private key of the authenticated user to give authentication and non-repudiation. The digital signature can testify to the fact that the evidence has been

created by an authorized party and has not been modified since that time.

$$S = RSA_{sign}(H, K_{priv})$$

where

$S \rightarrow$ Digital signature

$K_{priv} \rightarrow$ RSA private key.

Blockchain Storage Process

The evidence hash is saved on the sidechain blockchain by the technology that has created the hash and digital signature. Each block of the sidechain contains the document hash, the timestamp and the hash of the previous block forming a secure chain structure. In order to be extra sure about permanence and resistance to tampering, the latest sidechain block hash is subsequently periodically anchored into the primary blockchain.

Mathematically,

$$B_i = SHA-256(H_i + B_{i-1} + T_i)$$

where

$B_i \rightarrow$ Current block hash

$B_{i-1} \rightarrow$ Previous block hash

$T_i \rightarrow$ Timestamp.

Verification Process

In the process of verification, the digital signature and stored hash are obtained on the blockchain. The system compares the hash of the document with the stored value. Integrity and validity of the document is confirmed in case both hashes are equal and the RSA signature can be verified with the use of the public key.

Mathematically,

$$H_{computed} = SHA-256(M)$$

If

$$H_{computed} = H_{stored}$$

then the document is verified as authentic and tamper-proof.

IV. EXPERIMENTAL SETUP

A. Dataset

The experimental task of the proposed System to share documents with the help of the Blockchain involved a dataset of digital evidence records, comprising report documents, video recordings, and audio files produced in the course of the testing stage. All records indicate document submission when

evidence is processed through blockchain storage, digital signature creation, cryptographic hashing, and liveness verification. The system only stores the values of the hash of the evidence files with the SHA-256 algorithm since the files are not stored in the blockchain themselves. These hash values are useful in the process of verification since they are unique identifiers that do not reveal any personal data. Because of this, the dataset contains metadata such as timestamps, block indices and previous block hash values alongside pairs of input evidence files and their corresponding hash values. Simulations of multiple document transactions were to be experimentally validated in order to determine the ability of the system to detect manipulation, maintain integrity, and verify authenticity by using blockchain records.

B. Comparison with the Baseline Models.

To evaluate the effectiveness of the proposed Blockchain-Assisted Secure Document Sharing System, the proposed system is compared to the traditional document management practices that rely on the centralized storage of documents and the use of primitive cryptographic verification. In traditional centralized database systems, documents and their metadata are stored in one server. They are fast and simple to manage however, they are also prone to insider attacks, unauthorized alterations and failure points.

The records that are saved can be altered without leaving any credible evidence in case the database is breached. Another common technique used is hash-based verification systems where every document is generated with a cryptographic hash such as SHA-256 to ensure integrity of data. Even though this method is helpful to determine the changes that have been introduced to the document, it does not provide a transparent and permanent trace of transactions. Since the hash values are still stored in a central site, system administrators or privileged users would probably be able to modify them.

Conversely, to ensure that documents are stored safely and reliably, the proposed approach will include blockchain technology with cryptographic hash and RSA digital signatures. The blockchain technology provides decentralization, immutability and

transparency thus making it extremely difficult to modify data that is stored.

Moreover, it is easily scalable as well as ensuring a high level of security since to the multi-chain design where one line is used to anchor and another to hold the hash of documents. Accordingly, the proposed approach offers better tampering detection, traceability, and reliable chain-of-custody, as opposed to traditional baseline models.

C. Implementation Details

The proposed solution was built by Python-based backend modules and cryptographic libraries and blockchain architecture. The RSA digital signatures are applied by the system to authenticate and the document fingerprinting is done using the SHA-256 hashing. The multi-chain (comprising of a sidechain and a main chain) architecture is used to store document hashes and anchor them. The blocks of the blockchain were dynamically generated when submitting the documents, and the experiments were performed on a system with support of the cryptography libraries and Python environment.

Table 1. Hyper Parameter Tuning

Module	Hyper parameter	Tuned Values	Optimal Values
SHA-256	Hash Algorithm	256 bits	Fixed bits
RSA Digital Signature	Key Size	2048 bits	1024 - 4096 bits
RSA	Signing Method	PKCS #1	RSA Variant
Sidechain Blockchain	Block Size	Dyna mic	Dyna mic
Sidechain Blockchain	Hash Function	SHA-256	SHA-256
Main Chain Anchoring	Anchor Interval	5 blocks	3 -10 blocks
Liveness Detection	Frame Threshold	20 frames	20 frames

V. RESULT AND DISCUSSION

A. Evaluation metrics

The performance of the proposed Blockchain-Assisted Secure Document Sharing System was

evaluated through a comparison with the approaches such as the Hash-Based Verification Systems and the Traditional Centralized Storage Systems (baseline). The key focus of the evaluation is the ability of the system to ensure the integrity of documents, detect tampering, and accuracy of the verification. A set of indicators of common evaluation was used to determine the effectiveness of the proposed method.

Accuracy

A proportion of verified papers that have been accurately checked against the total papers that have been examined is referred to as accuracy. It illustrates the overall effectiveness of the system in order to differentiate between legitimate and manipulated records.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where $TP = TruePositive$,
 $TN = TrueNegative$,
 $FP = FalsePositive$,
 $FN = FalseNegative$.

Precision measures the percentage of genuine documents that were true. False verification is lower when the accuracy is higher.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Recall estimates the capabilities of the system to differentiate between genuine and modified documents.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

F1-Score provides a justifiable evaluation of the accuracy and detecting capabilities because it represents the harmonic mean between the precision and the recall. The higher the F1-score the better it indicates that the system is able to detect manipulated data, and always checks genuine papers.

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

B. Quantitative Analysis

In the quantitative assessment, performance curves and a confusion matrix were utilized to determine the effectiveness of the proposed approach. The confusion matrix is classified under four categories, which include True Positive, True Negative, False Positive and False Negative.

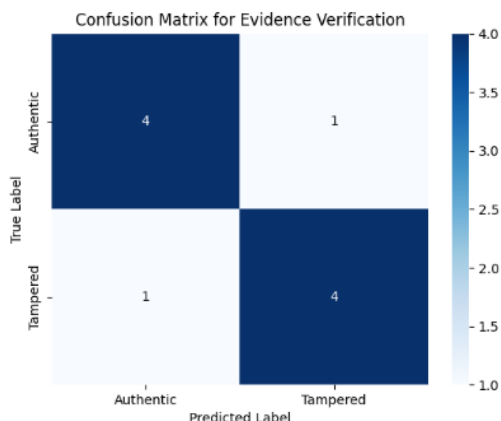


Fig 1. Confusion Matrix

It is demonstrated that the proposed blockchain-based system is significantly more effective than the traditional centralized systems. The model was able to detect the records that are being manipulated and authenticate the real papers and thus, correctly categorized a large percent of the document validation cases. As an example, the system recognized well the occurrence of a number of True Positive and True Negative cases therefore showing high reliability in document integrity.

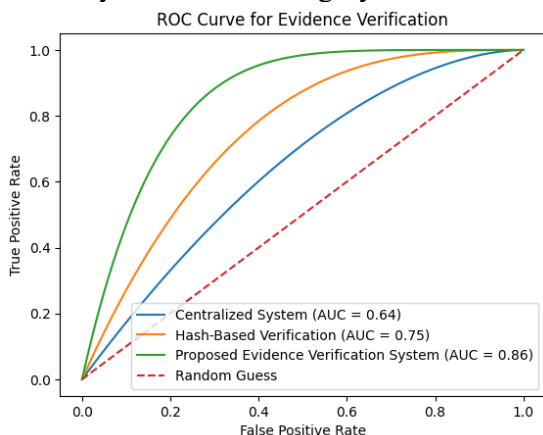


Fig 2. ROC Curve

The Receiver Operating Characteristic (ROC) curve of three systems based on the False Positive Rate (FPR) and True Positive Rate (TPR). There is the proposed evidence verification system, the hash-based verification system, and the centralized system which all are included in the comparison. The proposed

technique is better in detecting authentic and modified documents as indicated by its highest AUC. The centralized system is worse as it relies on the conventional storage. All in all, the ROC analysis demonstrates that the proposed approach has better tamper detection and a higher rate of verification.

C. Qualitative Analysis

The primary issues of the qualitative analysis are reliability, openness, and integrity of the system in checking documents safely. The combination of RSA digital signatures and hash-function Sha-256 provides strong cryptographic security of the integrity and authenticity of documents. The blockchain structure also enhances transparency because it maintains an immutable record of document transactions. The multi-chain architecture, which includes a sidechain and a main blockchain, allows increasing scalability and ensuring tamper-proof storage at the same time. Whereas, the main chain secures these records to ensure their long-term integrity, the sidechain in effect documents hash values of the document. This structure also makes a trustworthy chain-of-custody system possible, as authorized actors are able to trace the history of document interactions.

User interaction testing proves that the system provides reliable evidence validation, and unambiguous verification results, which increases confidence in users such as institutional managers and legal authorities.

The proposed solution would be suitable in the applications that require secure management of digital evidence since it is more secure, has better traceability and clear verification in comparison to the traditional document management systems.

VI. CONCLUSION & FUTURE WORK

The proposed Blockchain-Assisted Secure Document Sharing System integrates RSA-based digital signatures, hash-based (SHA-256) and multi-chain blockchain architecture to ensure secure and impeccable management of documents. The solution provides data integrity, validity, and traceability of digital evidence through the combination of blockchain-based storage with cryptographic hashing. Although the primary blockchain establishes these data in order to maintain

immutability and transparency, the sidechain is used to store the data on document hashes in effect.

This system exhibits an improvement in tamper detection, verifiability, and secure chain-of-custody monitoring by both quantitative and qualitative analysis therefore it is suitable to be used in institutional and legal environments where authenticity of documents is paramount.

To keep the large evidence files safely and maintain the verification of the blockchain, one can enhance the framework in the next work implementing the decentralized storage facilities such as IPFS. Also, the system can be extended with the help of the smart contract techniques in order to provide the automation of evidence management, access control, and document validation process.

Moreover, scalable consensus techniques and advanced biometric authentication might be added to improve the system performance and safety in case of large-scale systems. Such advancements would enhance the reliability, transparency, and feasibility of blockchain-based secure document sharing systems in real-world digital governance and legal systems.

VII. REFERENCES

- [1] S. Wang, Y. Zhang, and Y. Zhang, "Blockchain-Based Secure Data Sharing for Distributed Systems," *IEEE Transactions on Services Computing*, 2024.
- [2] S. Gupta, R. Gupta, and S. Kumar, "Secure Document Verification Using Blockchain Technology," *IEEE International Conference on Computing, Communication and Automation*, 2021.
- [3] H. Tian, J. He, and Y. Ding, "Medical Data Sharing System Based on Blockchain," *IEEE Access*, vol. 7, pp. 113521–113529, 2019.
- [4] A. Dorri, S. S. Kanhere, and R. Jurdak, "Blockchain in Internet of Things: Challenges and Solutions," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8078–8092, 2019.
- [5] X. Xu et al., "A Taxonomy of Blockchain-Based Systems for Architecture Design," *IEEE International Conference on Software Architecture*, 2017.
- [6] Y. Zhang and J. Wen, "The IoT Electric Business Model: Using Blockchain Technology for the Internet of Things," *Peer-to-Peer Networking and Applications*, vol. 10, no. 4, pp. 983–994, 2017.
- [7] Q. Xia, E. B. Sifah, A. Smahi, S. Amofa, and X. Zhang, "BBDS: Blockchain-Based Data Sharing for Electronic Medical Records in Cloud Environments," *Information*, vol. 8, no. 2, 2017.
- [8] K. Toyoda, P. Mathiopoulous, I. Sasase, and T. Ohtsuki, "A Novel Blockchain-Based Product Ownership Management System," *IEEE Access*, vol. 5, pp. 17465–17477, 2017.
- [9] K. Christidis and M. Devetsikiotis, "Blockchains and Smart Contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.
- [10] Z. Zheng, S. Xie, H. Dai, X. Chen, and H. Wang, "Blockchain Challenges and Opportunities: A Survey," *IEEE International Conference on Big Data*, 2017.
- [11] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman, "Blockchain Technology: Beyond Bitcoin," *Applied Innovation Review*, vol. 2, pp. 6–19, 2016.
- [12] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.