Blockchain for Privacy-Preserving Data Distribution in Healthcare

Amitesh Singh Rajput and Arnav Agrawal
Department of Computer Science & Information Systems, Birla Institute of Technology & Science, Pilani, Vidya Vihar, Pilani, Rajasthan, India

Keywords: Blockchain Technology, Healthcare Data Security, Decentralized Data Management, Smart Contracts in Healthcare, Encryption and Key Management, Authentication Protocols.

Abstract: As virtual transformation maintains to reshape healthcare, the security and privacy of health information have become paramount worries. This paper delves into the novel application of blockchain generation as a strategic technique to these urgent issues. In contrast to traditional centralized information control structures, blockchain introduces an intensive alternate with its decentralized, immutable, and transparent nature. This shift gives a robust alternative to comply with sensitive health data. We propose a contemporary, blockchain-primarily based method to seamlessly integrate existing healthcare records into ledgers and share them in a controlled way. The proposed method emphasizes enhanced data integrity, advanced security features, and a patient-centric technique to data governance using customized smart contracts. Experimental results underline the proposed method's advanced performance for scalability, protection, and general machine performance, making a compelling case for its adoption in healthcare records control.

1 INTRODUCTION

In today's complex digital world, the healthcare sector is particularly sensitive due to its critical nature. With an unparalleled emphasis on the sanctity, safety, and accessibility of facts, healthcare gives multifaceted challenges. The very lifeblood of medical practices and studies, patient records, is often saved in centralized databases. While these centralized structures have long been taken into consideration as reliable workhorses, they are being more identified for their inherent flaws and vulnerabilities (Aslan et al., 2023).

Centralized healthcare databases, although efficient in many respects, are susceptible to a variety of challenges. These range from system failures, cyber-attacks, unauthorized data breaches to potential data tampering (Newaz et al., 2021). Furthermore, the modern demand for seamless data flow and integrated healthcare operations often stumbles against the brick walls of these fragmented and isolated systems. Our study addresses this problem by providing a blockchain-based model for healthcare data distribution. Blockchain technology offers a promising answer to overcome the limitations of conventional centralized systems (Haleem et al., 2021). In the healthcare sector, where data sensitivity and privacy are paramount, blockchain can revolutionize how person's data is stored, accessed, and shared. A blockchain-based system can mitigate risks associated with centralized databases by means of dispensing the data across a network, making it less vulnerable to single points of failure and cyber assaults. This guarantees better data integrity and protection. Additionally, blockchain's transparent nature allows better trust and auditability in data transactions, that is critical in healthcare wherein data accuracy and history are important.

In recent years, several blockchain-based medical recordkeeping systems have been developed, offering innovative solutions to the challenges in healthcare data management. For instance, MedRec (Azaria et al., 2016a) is a notable system that uses Ethereum blockchain for managing access to medical records. It provides a decentralized record management system to handle authentication, confidentiality, accountability, and data sharing. HealthChain (Chenthara et al., 2020) focuses on ensuring data integrity and access control in healthcare records. It allows for a secure and transparent record-keeping mechanism that patients and providers can trust. MediBlock (Vishwa, 2021) is another significant advancement in this space, providing a patient-centered approach. It emphasizes patient consent and uses blockchain to secure medical records.

Key Authorities (KAs) are essential in blockchain-based healthcare data systems. They

Rajput, A. and Agrawal, A.
Blockchain for Privacy-Preserving Data Distribution in Healthcare.
DOI: 10.5220/0012470500003648

In *Proceedings of the 10th International Conference on Information Systems Security and Privacy (ICISSP 2024)*, pages 621-631
ISBN: 978-989-758-683-5; ISSN: 2184-4356

are accountable for dealing with cryptographic keys, making sure that the data is secure, maintaining data integrity, facilitating secure transactions, and contributing to system scalability. KAs play an important role in protecting sensitive healthcare data, making them essential to the security and capability of such systems. Despite the innovative and potential solutions for the health records, previous researches (Azaria et al., 2016a)(Chenthara et al., 2020)(Vishwa, 2021) have not emphasized on this fundamental part. We consider this part as utmost important along with robust patient data sharing and propose a blockchain based automated approach in this paper. We propose a comprehensive set of rules, intricately designed, showcasing how blockchain can be seamlessly integrated into the healthcare facts pipeline involving an extensive functioning for KAs. By examining its core attributes, we unravel how blockchain could be the keystone in constructing a robust healthcare data ecosystem.

This paper aspires to make a case for a blockchain-driven healthcare data revolution. In the following sections, we will discover a meticulous breakdown of current challenges in section 2, our proposed blockchain solution and its practical algorithmic descriptions in section 3, followed by a thorough experimental analysis and comparison with the existing blockchain methods in section 4. We believe that our research not only contributes extensively to academic discourse but additionally paves the manner for practical, tangible upgrades in the real world of healthcare data management.

2 LITERATURE REVIEW

Data distribution, especially in an age of escalating cybersecurity threats and stringent privacy regulations, is a domain of paramount importance. Addressing countless challenges in this area necessitates a holistic understanding of both the historical context and the forefront of current advancements. This literature review describes the landscape of data distribution, particularly emphasizing the following three pivotal facets:

1. **Challenges of Traditional Data Distribution Systems.** An exploration into the pitfalls and vulnerabilities of legacy systems, highlighting the need for innovative solutions.
2. **Blockchain's Emergence in Privacy-Preserving Data Distribution.** Recognising blockchain's revolutionary potential to reimagine safe and transparent data distribution paradigms.
3. **Existing Access Control Methods Integrating Multiple Entities in the Overall System.** A thorough examination of prior methods combining several entities to overcome current obstacles.

2.1 Challenges of Traditional Data Distribution Systems

Traditional data distribution systems play an integral role in various sectors, especially in the age of digital transformation. However, their inherent design and operational characteristics present multiple challenges.

Centralized data distribution boasts ease of management, uniformity of data versions, and straightforward administrative control (Inclusion Cloud, 2023). An attack on this system can render the entire data distribution network vulnerable resulting in irreversible data loss (Mohammed et al., 2010).The linear scaling model of infrastructure leads to longer response times. In cases even system downtime. This negatively impacts both user experience and operational efficiency (Bertin et al., 2009). Centralized data servers are typically located in different regions. Consequently they become subject to laws that may not align with operations or could impose restrictions on data access. This can lead to inconsistencies in operations (Salmon and Myers, 2019).With all data stored in a location there is a risk of monopolization where the controlling entity may misuse the information (McIntosh, 2018). Without immutable record-keeping, changes to data can go undetected. This lack of an unalterable history can result in large commercial enterprise disputes, mainly in sectors wherein data sanctity is paramount (Pandey et al., 2020). Traditional systems might not always provide mechanisms for real-time data validation, making it difficult to immediately detect and rectify data inconsistencies (Kumar and Bhatia, 2020).

The centralized model, in many instances, does not maintain transparent data modification logs. Such opacity makes it challenging to trace data changes, leading to potential trust breaches (Abiteboul and Stoyanovich, 2019). All decisions, including data access, distribution rules, and dispute resolutions, are at the discretion of a central authority. Such a system might not always be in the best interests of all stakeholders involved (Laoutaris, 2018). The centralized nature necessitates regular maintenance, leading to recurrent costs and potential downtimes (Chukmaitov et al., 2015). They consume enormous amounts of energy, leading to operational inefficiencies and larger carbon footprints (Chukmaitov et al., 2015). Centralized servers, especially during peak times, can face

bandwidth constraints, slowing down data dissemination rates (Hugoson, 2009).

Researchers and developers around the world have been investigating alternate paradigms for data delivery in light of these difficulties. With its decentralised structure, immutable record-keeping, and strong security features, blockchain technology has gained attention as a possible solution to several of these issues (Bhutta et al., 2021).

2.2 Blockchain's Emergence in Privacy-Preserving Data Distribution

Blockchain technology stands out as a highly promising option for data distribution including built-in attributes like traceability, transparency, and resistance to tampering (Attaran, 2022). Numerous approaches that take advantage of blockchain technology have been put out and investigated; these have primarily concerned healthcare and distribution while maintaining anonymity. Here, we examine some of the most significant developments in this field, shedding light on novel approaches and their consequences.

2.2.1 Blockchain in Healthcare: A Paradigm Shift

Blockchain, at the start conceptualized for the Bitcoin cryptocurrency (Nakamoto, 2008), has found great applications in healthcare, especially in protecting the integrity, authenticity, and confidentiality of health records (Azaria et al., 2016b). The decentralized and immutable nature of the blockchain ensures that every transaction, or in the case of healthcare, every data entry or access, is transparently recorded and verified by a network of peers.

MedRec, introduced by Azaria et al. (Azaria et al., 2016b), is an influential work that offers a decentralized electronic health records system using Ethereum's blockchain. This system presents solutions to the challenges of data interoperability, access controls, and auditability in traditional EHR systems. By granting patients the power to grant access to their health records and ensuring a transparent record of data access, MedRec paves the way for improved patient-centric healthcare.

Furthermore, Yue et al. (Yue et al., 2016) proposed a scheme named "Healthcare Data Gateways", leveraging blockchain to ensure data security, privacy, and interoperability in health information exchange. The framework emphasizes the role of patient consent in data sharing and makes use of smart contracts to automate and enforce data access policies.

2.2.2 Privacy-Preserving Distribution with Blockchain

While healthcare is a predominant domain, the extensive theme of privacy preservation facilitated by blockchain extends to various data distribution models. Zhang et al. (Zhang and Wen, 2017) introduced a privacy-preserving framework which utilizes blockchain for secure and transparent file access in cloud environments. This framework harnesses the decentralized consensus mechanism of blockchain to deter unauthorized data access and modifications.

Another notable advancement is the work of Liang et al. (Liang et al., 2017), where they propose a credit-based scheme for privacy-preserving data sharing in the context of the Internet of Things (IoT). Their methodology combines the blockchain with a credit system, ensuring that data providers and consumers maintain a trust-based relationship, with transactions transparently recorded and verified on the blockchain.

In addressing the challenge of non-interoperable medical data sharing, there is a blockchain-based system employing symmetric encryption, access control, and participant power restriction. Their system ensures privacy by encrypting medical data, utilizing a novel blockchain framework, implementing chameleon signatures, and enabling revocable participant privileges. Security analysis validates its robustness, and experiments demonstrate superior time overhead compared to alternatives (Hu et al., 2023).

In response to security challenges in medical data transmission and sharing, Chen proposed a blockchain-based medical information system ensuring data integrity and privacy. Their model employs IoT for real-time patient health record collection and introduces a secure, anonymous data sharing scheme based on cloud servers and proxy re-encryption. Implemented on Hyperledger Fabric, the system features a dual-channel architecture and medical chaincode for efficient data management and access control. (Chen et al., 2021).

Huang presented a blockchain-based privacy-preserving scheme addressing the challenge of balancing patient privacy with the demands of health data research and commerce. The proposed scheme enables secure sharing of medical data among entities, including patients, research institutions, and semi-trusted cloud servers. Utilizing zero-knowledge proof for privacy-preserving verification and proxy re-encryption for intermediary data decryption, the scheme ensures data availability and consistency. (Huang et al., 2020).

2.3 Existing Access Control Methods Integrating Multiple Entities in the Overall System

The landscape of data distribution has seen countless techniques aimed at ensuring robustness, efficiency, and security. A few of the previous methods (Hildebrandt et al., 2016)(Erkin et al.,)(Rajput and Balasubramanian, 2021) heavily rely with an integration of multiple entities including KAs, transcryptors (TRs), doctors, storage facilities, and medical researchers. The TRs serves as intermediary entities. When a request for data access is made, it is the TR's role to fetch the corresponding cryptographic key from the KA, perform the necessary decryption operations, and furnish the requester with the requisite decrypted data. Other entities may encompass data providers, end-users, or even audit entities ensuring the system's integrity and adherence to established protocols. The functioning is performed with the data provider encrypting the data using cryptographic keys sourced from the KAs. Once encrypted, the data is set for distribution. Upon receipt of a valid data access request, the TR gets into action. It fetches the associated cryptographic key, decrypts the data, and ensures its timely delivery to the requester.

One strong challenge with these methods is their reliance on centralized entities, i.e., the KAs. Centralization, while offering streamlined operations can be a vulnerability concern. Compromising the KA could potentially jeopardize the entire system's data security. Latency is another big concern which needs to be taken care of. Given the multiple interactions between entities, especially between TR and KAs for key retrieval, time-sensitive operations can be adversely affected. While the algorithm integrating KAs, TRs, and other entities showcased potential in addressing data distribution's challenges, it wasn't devoid of limitations. The centralization concerns, latency issues, and scalability bottlenecks underlined the need for alternative mechanisms that could offer both efficiency and robust security. It is within this context that the exploration for newer methods, including those harnessing blockchain, emerged.

3 PROPOSED METHOD

This section describes the proposed blockchain-based method, designed to triumph over the restrictions of traditional data distribution systems, particularly inside the healthcare sector. At its center, our approach leverages the decentralized, transparent, and immutable nature of blockchain generation. The approach is multifaceted, integrating advanced cryptographic strategies, smart contracts, and consensus algorithms to enhance data integrity, security, and transparency. We aim to decentralize data management for getting rid of single points of failure, improving scalability, and increasing resistance to cyber threats. Furthermore, the proposed approach introduces a patient-centric version for data governance, ensuring privacy and hold over personal health records.

In the subsequent subsections, the critical components of the proposed approach are described. Also, the setup of blockchain-based identity management, secure packet sharing mechanisms, and the processes for master key generation, data request handling, and user verification are covered. These additives collectively form a cohesive and robust framework for healthcare data management, ready to be incorporated into current healthcare IT infrastructures. The description contains various specialized algorithms important with respect to the proposed solution. It is important to note that the functions in these algorithms are named to clearly indicate their purpose. This naming approach helps in easily understanding what each function does and how it fits into the overall system.

In light of the limitations encountered with the traditional data distribution techniques and the inefficacies linked to the previously discussed algorithm integrating KAs, TRs, and other entities, the proposed proposition pivots toward leveraging the capabilities of blockchain technology. The emphasis here is not only to decentralize the data distribution process but also to harness the inherent security and transparency features of blockchain. The proposed method leverages customized smart contracts that play a pivotal role to accomplish the desired tasks. Essentially, smart contracts are self-executing contracts where the agreement between the sender and the receiver is directly written into lines of code. In the context of our data distribution system, the smart contracts will automate the validation of data access requests and the distribution of decryption keys.

In the proposed method, smart contracts execute algorithms linearly for optimal functionality. The sequence begins with *Identity Registration and Blockchain Logging* (Algorithm 1), establishing verified user identities on the blockchain. Next, the *Random Selection of Key Authorities* (Algorithm 2) ensures unbiased selection of Key Authorities. This is followed by *Secure Verification Packet Sharing* (Algorithm 3), distributing verification data to KAs. Subsequently, *Master Key Generation and Validation* (Algorithm 4) securely generates encryption keys, and *Distribution of Verification Packets* (Algorithm 5) ensures secure data dissemination to KAs. *Data*

Request, Verification & Blockchain Logging (Algorithm 6) then manages and authenticates data requests. *Master Key Derivation & Decryption Process* (Algorithm 7) and *Data Decryption & Sharing* (Algorithm 8) handle the secure decryption and distribution of data. The process concludes with *Repetitive User Verification* (Algorithm 9), maintaining continuous security checks. This ordered execution of algorithms as smart contracts solidifies the system's integrity and efficiency in healthcare data management.

3.1 Setting up an Identity and Registering on the Blockchain

In this section, the algorithm describes the way of organising a person's identification inside the system and securing it via blockchain technology. The method begins with the generation of a completely unique identifier, accompanied by the introduction of a digital credential which is then stored in a database for quick access. To ensure the individuality and prevent duplication of identities, the credential is verified with the prevailing facts on the blockchain. If the credential is new, it is registered at the blockchain; otherwise, a duplication error is reported. This approach as present in the algorithm 1 ensures that every person within the machine has a distinct and verifiable digital identification.

Algorithm 1: Identity Registration and Blockchain Logging.

Input: User's name *name*, blockchain *BC*
Output: Registration status
uniqueID ← GENERATEUNIQUEID(*name*)
credential ← (*name*, *uniqueID*)
STORECREDENTIALS(*credential*)
if ¬VERIFYIDENTITY(*credential*, *BC*) **then**
    REGISTERONBLOCKCHAIN(*credential*, *BC*)
else
    print "Error: Duplicate identity."
end if
function VERIFYIDENTITY(*credential*, *BC*)
    for all *block* ∈ *BC* **do**
        if *block.data* = *credential* **then**
            return true
        end if
    end for
    return false
end function
function STORECREDENTIALS(*credential*)
    // Store credentials in the database
end function
function REGISTERONBLOCKCHAIN(*credential*, *BC*)
    // Add new identity to the blockchain
end function

3.2 Secure Verification Packet Sharing

This section establishes a procedure for the secure creation and dissemination of verification packets, crucial for the operations of Key Authorities within the blockchain network. It details a series of steps, mentioned in algorithm 2, that begin with the generation of a verification packet, which is uniquely identified by a hash and timestamp to ensure its integrity. These packets are then stored securely until they are to be shared with eligible Key Authorities, with each sharing event logged on the blockchain to enforce transparency and traceability.

3.3 Random Selection by TR and Recording on Blockchain

The selection steps in algorithm 3 operates with a dual aim: To guarantee fairness in the selection process of KAs and to log each selection event onto the blockchain, thus embedding transparency and immutability into the system. It begins by identifying KAs that have not been recently active, thereby ensuring an unbiased opportunity for all. A rigorous safety check follows, verifying the temporal gap since each KA's last selection, thus preventing repetitive selections. Successful candidates are then recorded on the blockchain, time-stamped to uphold the integrity and verifiability of the process.

Algorithm 2: Random Selection of KA with Block Addition.

Input: Pool of Key Authorities *KA_Pool*, blockchain *BC*
Output: Selected Key Authority *selectedKA*
function SELECTKEYAUTHORITY(*KA_Pool*)
    selectedKA ← Select KA not recently chosen
    if SAFETYCHECK(*selectedKA*) **then**
        RECORDSELECTION(*selectedKA*, *BC*)
    else
        print "Selection failed: KA chosen recently."
    end if
end function
function SAFETYCHECK(*selectedKA*)
    $timeDiff \leftarrow$ Time since *selectedKA*'s last selection
    return $timeDiff \geq 24$ hours
end function
function RECORDSELECTION(*selectedKA*, *BC*)
    ADDBLOCK(*BC*, "SELECTION", *selectedKA*, "current time")
end function

Algorithm 3: Secure Verification Packet Sharing.

Input: Data for Verification Packet *data*, Key Authorities *KA*
Output: Status of Verification Packet Sharing
function CREATEANDSHAREVERPACKET(*data*, *KA*)
    packet ← CREATEPACKET(*data*)
    for all *ka* ∈ *KA* **do**
        if ISELIGIBLE(*ka*) **then**
            SHAREPACKET(*packet*, *ka*)
            LOGPACKETSHARING(*packet*, *ka*)
        end if
    end for
end function

3.4 Master Key Generation with Blockchain Validation

The algorithm 4 outlines a secure process for Master Key (MK) generation, encryption, and storage, coupled with blockchain logging for validation. The MK, generated with a secure random function and timestamped, is encrypted for confidentiality. It's then stored in a secure database, with blockchain logging ensuring transparency and validation checks maintaining generation at safe intervals.

3.5 Distribution of Verification Packets

The secure and efficient distribution of Verification Packets (VP) to Key Authorities (KA) within a blockchain network is the basis of this system. The algorithm initiates with the generation of a unique packet ID and its subsequent encryption. KAs are selected based on specific criteria for packet distribution, with the entire process being logged onto the blockchain for accountability. Duplicate distributions are prevented through a verification mechanism, ensuring the integrity of the system. All this can be verified from the Algorithm 5.

Algorithm 4: Master Key Generation and Validation.

Input: Security parameters, blockchain *BC*
Output: Master Key Generation and Validation
function GENERATEANDVALIDATEMASTERKEY
    masterKey ← GENERATEMASTERKEY
    encrKey ← ENCRMASTERKEY(*masterKey*)
    STOREENCRYPTEDKEY(*encrKey*)
    if VALIDATEKEYGENERATION **then**
        LOGKEYGENERATION(*encrKey*, *BC*)
    else
        print "Key generation validation failed."
    end if
end function

3.6 Data Request, Verification & Blockchain Logging

This subsection outlines the streamlined process of receiving data requests, verifying requester authenticity, and logging the transactions on the blockchain. The Algorithm 6 handles requests with a security-focused approach, ensuring data integrity and access control.

3.7 Master Key Derivation & Decryption Process

The Algorithm 7 below starts with the derivation of a master key (MK) which is an important step in ensuring the security of the system. It is created using a secure function and is timestamped to guarantee uniqueness. After derivation, the MK is encrypted and stored securely. It is then utilized for decrypting data, enforcing confidentiality. The system logs each step involving the MK on the blockchain.

Algorithm 5: Distribution of Verification Packets.

Input: Verification Packet data *data*, Key Authorities *KA*
Output: Distribution status
function DISTRIBUTEVERPACKET(*data*, *KA*)
    packetID ← GENERATEPACKETID(*data*)
    encryptedPacket ← ENCRYPTPACKET(*data*)
    selectedKAs ← SELECTKASFORDISTRIBUTION
    for *ka* in *selectedKAs* **do**
        LOGDISTRIBUTIONEVENT(*packetID*, *ka*)
        DISTRIBUTEPACKET(*encryptedPacket*, *ka*)
        print "Packet distributed to KA"
    end for
    CONFIRMFROMKA(packetID, selectedKAs)
end function

Algorithm 6: Data Request, Verification & Blockchain Logging.

Input: Request from entity *requester*, data identifier *dataID*
Output: Acknowledgment of successful data handling
function HANDLEDATAREQUEST(*requester*, *dataID*, *BC*)
    isAuthentic ← VERIFYREQUESTER(*requester*)
    if *isAuthentic* & SAFETYCHECK(*requester*) **then**
        data ← FETCHDATA(*dataID*)
        encryptedData ← ENCRYPTDATA(*data*)
        TRANSMITDATA(*encryptedData*, *requester*)
        LOGTRANSACTION(*requester*, *dataID*, *BC*)
        return AWAITRECEIPT(*requester*)
    else
        print "Request denied."
        return False
    end if
end function

3.8 Data Decryption & Sharing

The secure distribution of Verification Packets is a critical characteristic in our blockchain-based system, in particular for Key Authorities (KA) who rely upon these packets for operational data. Each packet is created with a completely unique identifier, encapsulated securely, after which it is transmitted to a selected group of KAs. This method as mentioned in the Algorithm 8 guarantees information integrity through encryption and maintains transparency by recording each distribution motion at the blockchain.

Algorithm 7: Master Key Derivation and Decryption.

Input: *inputSeed*, *encryptedData*
Output: *decryptedData*
function DERIVEANDENCRYPTMK(*inputSeed*)
    masterKey ← derive key using *inputSeed*
    encryptedMK ← encrypt *masterKey*
    Store *encryptedMK* securely
    Log MK creation on blockchain
    return *encryptedMK*
end function
function DECRYPTWITHMK(*encryptedData*, *masterK*)
    decryptedData ← decrypt *encryptedData* with *masterKey*
    Log decryption on blockchain
    return *decryptedData*
end function

Algorithm 8: Distribution of Verification Packets.

Input: Verification Packet data *data*, Key Authorities *KAs*
Output: Distribution status
function DISTRIBUTEVERIFICATIONPACKET(*data*, *KAs*)
    packetID ← GENERATEPACKETID(*data*)
    encryptedPacket ← ENCRYPTPACKET(*data*)
    selectedKAs ← SELECTKASFORDISTRIBUTION
    for all *ka* ∈ *selectedKAs* **do**
        if SAFETYCHECK(*ka*) **then**
            SENDPACKET(*encryptedPacket*, *ka*)
            LOGPACKETSHARING(*packetID*, *ka*)
            print "Packet distributed to KA: *ka*"
        else
            print "Distribution to KA: *ka* failed due to safety check."
        end if
    end for
    return CONFIRMRECEIPT(*selectedKAs*)
end function

3.9 Repetitive User Verification

The Repetitive User Verification process mentioned in the Algorithm 9 within the system is designed to set up and preserve the authenticity of user identities. By capturing user details and validating identities, the device ensures that only legitimate users can get entry to sensitive information. The blockchain logs each verification attempt, providing an immutable record of user activity. Enhanced safety features, along with multi-factor authentication and checks for fraudulent tries, shield against unauthorized access. Any unusual activities are monitored and reported, underscoring the system's commitment to security and user trust.

Algorithm 9: Repetitive User Verification.

Input: User ID *userID*, Verification Details *userDetails*
Output: Verification Status
function VERIFYUSER(*userID*, *userDetails*)
    details ← CAPTUREUSERDETAILS(*userDetails*)
    if VALIDATEUSERIDENTITY(*details*) **then**
        LOGVERIFICATIONEVENT(*userID*, success)
        INITIATEMULTIFACTORAUTHENTICATION(*userID*)
        print "User verified successfully."
    else
        if SAFETYCHECKFORFRAUDULENTATTEMPTS(*userID*) **then**
            LIMITUNSUCCESSFULATTEMPTS(*userID*)
            REPORTSUSPICIOUSACTIVITY(*userID*)
            print "User verification failed."
        end if
    end if
end function

4 EXPERIMENTAL RESULTS

The experimental results presented here are derived from simulations that model the overall performance of the proposed blockchain-based system. Our comprehensive evaluation specializes in several key performance indicators inclusive of throughput, latency, scalability, security, and smart contract execution metrics. The simulations have been performed on a virtualized test network replicating a medium-scale enterprise environment to offer a realistic evaluation of system capabilities.

4.1 Simulation Environment

The testing environment set up specifications include Ubuntu 20.04 LTS Operating System, Intel Xeon CPU E5-2698 v4 @2.20GHz Processor, 64GB RAM,

1 TB SSD, Ethereum Testnet blockchain platform, followed by Proof of Authority (PoA) consensus algorithm, and 25 validator nodes.

The selection of Proof of Authority (PoA) as the consensus mechanism is rooted in its suitability for healthcare data management. PoA provides a secure and efficient environment, crucial for handling sensitive health records. Key reasons for choosing PoA include:

- **Trust and Identity Verification.** PoA ensures validators are known entities, enhancing trust and security, essential in healthcare for regulatory compliance.
- **Efficient Transaction Processing.** PoA offers faster transaction times compared to PoW or PoS, crucial for real-time healthcare applications.
- **Energy Efficiency.** Less energy-intensive than PoW, PoA aligns with sustainable IT practices in healthcare.
- **Controlled Network Access.** Restricts network participation to verified entities, maintaining data privacy and adhering to healthcare regulations.
- **Balance of Centralization.** PoA's selective validator approach mitigates risks associated with centralization, ensuring decentralized control.

4.2 Throughput and Latency

The table showcasing throughput and latency illustrates how the system effectively handles growing numbers of transactions. As the transaction count rises from 100 to 5000, the system's throughput, measured in Transactions Per Second (TPS), increases. However, past a certain point (at 5000 transactions), a slight descrease in throughput is found, probably because of the increased load. Latency, or the time taken to complete a transaction, understandably increases with the range of transactions, but stays inside a reasonable range, showing the system's responsiveness even under heavy load.

Table 1: System Throughput and Latency.

Number of Transactions	Throughput (TPS)	Latency (Seconds)
100	50	2
500	150	4
1000	200	6
5000	180	10

4.3 Scalability Analysis

The Scalability analysis table reflects the system's ability to maintain overall performance as the nodes increase in number. The average transaction time shows a slow growth as extra nodes are brought, indicating a modest effect on performance. This upward push is highly linear and mild, suggesting that the device scales well. The capacity to deal with extra nodes with only a minor increase in transaction time is a strong indicator of the system's scalability and its capacity for larger-scale deployment.

Table 2: Scalability Test Results.

Number of Nodes	Average Transaction Time (Seconds)
10	2
20	2.2
30	2.5
40	2.75
50	3

4.4 Security Analysis

The security evaluation was executed to evaluate the resilience of the proposed blockchain-based system against various attack vectors, taking into account its unique structure and functionalities. The evaluation covered exclusive threat models which can be relevant to the system's operational context in healthcare data management.

4.4.1 Threat Models and Resistance Mechanisms

1. **DDoS Attack.** The system's decentralized nature, combined with PoA consensus, inherently resists DDoS attacks, as attackers cannot target a single central point. The distributed validator nodes ensure continuous network functionality even under high load.

   Proof. The blockchain system employs a distributed network of validator nodes. Let N be the total number of nodes in the network. The probability P of successfully disrupting the network is given by $P = \left(\frac{1}{N}\right)^n$, where n is the number of nodes an attacker needs to compromise. Given the large number of nodes in a distributed system, this probability approaches zero, thus demonstrating the resilience of the system against DDoS attacks.

2. **Double-Spending.** The blockchain's immutable ledger prevents double-spending by maintaining a transparent and unalterable transaction his-

tory. Each transaction is verified and recorded across multiple nodes, making fraudulent attempts highly impractical.

Proof. Assume a transaction T is initiated. For T to be double-spent, it must be replicated as T'. However, the blockchain's consensus mechanism requires that all transactions be verified by multiple nodes. The probability of T' being accepted by all verifying nodes without detection is negligibly small. Hence, the system effectively mitigates the risk of double-spending.

3. **Sybil Attack.** The PoA mechanism requires validators to be pre-approved, limiting the potential for Sybil attacks. Validators are trusted entities, reducing the risk of malicious nodes entering the network.

   Proof. Consider a network with a set of pre-approved validators using the PoA consensus. Let V be the set of all validators and S be the subset of malicious validators an attacker attempts to introduce. The system's trust model and validator approval process ensure that $|S| \ll |V|$ that reduces the effectiveness of any Sybil attack.

4. **Insider Threats.** Smart contracts use for critical operations such as data access and identity verification minimizes human intervention, thereby reducing the risk of insider threats.

   Proof. Smart contracts in the system autonomously execute predefined tasks without human intervention. Let I be the set of all potential insider attackers. The absence of direct control over data by any member of I significantly reduces the probability of successful insider threats, ensuring the security of the system against such attacks.

Each threat model was addressed through a combination of blockchain's inherent features and specific design choices made for the system. While analysing security breach attempts, different attack types were rigorously examined against our systems. There were approximately 1000 attempts of Distributed Denial of Service (DDoS) attacks, 500 attempts of Double-Spending attacks, 300 Sybil attacks, and 200 instances of Insider Threats. These type of attempts resulted in a 0% breach success rate, demonstrating the robustness and effectiveness of our security measures.

4.5 Smart Contract Performance

Table 3 provides insights into the performance of key smart contracts inside the system. It gives information regarding the execution time and gas fee for every agreement, such as Identity Registration, Data Request, and Verification Packet. The execution times are brief, demonstrating the system's performance in processing these critical operations. The gas costs, a measure of the computational resources required, are notably low, indicating an economically feasible model for running these contracts on the blockchain network.

Table 3: Smart Contract Execution Metrics.

Smart Contract	Execution Time (ms)	Gas Cost
Identity Registration	150	0.0003
Data Request	200	0.0004
Verification Packet	250	0.0002

4.6 Discussion and Comparative Analysis with MedRec

Experimental results validate the technical and practical viability of the proposed blockchain-based system. It successfully handles a huge quantity of transactions with minimal latency, an important feature for healthcare applications requiring real-time data processing. The scalability evaluations show that the system sustains performance despite increased network size, appropriate for tremendous deployment.

The system demonstrates strong resilience towards common cyber threats like DDoS attacks, double-spending, Sybil attack and Insider threats, underlining its robustness for healthcare data management wherein safety and patient privacy are essential. Smart contracts, pivotal for automating key processes along with identification verification and data requests, perform successfully in terms of execution time and useful resource usage. The experimental outcomes suggest the machine's readiness for healthcare implementation. Its extremely good throughput, low latency, protection and scalability, coupled with smart contract performance, gift it as a promising answer for transforming healthcare data control.

A comparative analysis between the existing blockchain based healthcare data management method MedRec (Azaria et al., 2016a) and the proposed method is presented in Table 4. This comprehensive assessment aims to highlight the distinctions and similarities in both the approaches to data protection, patient consent, interoperability, and other important aspects of healthcare data management.

The comparative analysis sheds light on the respective strengths and improvement areas of both systems. While the MedRec system lays a foundational framework using blockchain for healthcare data, our proposed approach builds and expands upon these

Table 4: Comparison of proposed method and MedRec (Azaria et al., 2016a).

Feature	**Proposed method**	**MedRec (Azaria et al., 2016a)**
Blockchain Platform	Utilizes a custom blockchain architecture.	Employs Ethereum blockchain.
Data Security	Enhanced through advanced cryptographic techniques.	Relies on Ethereum's inherent security features.
Patient Consent	Automated consent via smart contracts.	Patient consent managed through Ethereum transactions.
Interoperability	Designed for seamless integration with existing healthcare systems.	Focuses on interoperability within the Ethereum ecosystem.
Data Access Control	Controlled by smart contracts for precise data governance.	Access control is managed through Ethereum smart contracts.
Scalability	Addresses scalability with a tailored approach to healthcare data.	Faces typical scalability challenges of Ethereum.
Consensus Mechanism	Uses Proof of Authority (PoA) for efficiency and security.	Dependent on Ethereum's consensus mechanism.
Patient Data Privacy	Prioritizes patient privacy with decentralized data management.	Ensures privacy through blockchain's transparent nature.
Audit Trails	Transparent and immutable audit trails for data transactions.	Similar approach using Ethereum's ledger.
Customizability	High degree of customization for healthcare needs.	Limited by Ethereum's platform constraints.

concepts. By introducing enhanced scalability, customizability, and better integration capabilities, the proposed method demonstrates considerable advancements over the MedRec system, further advancing blockchain technology's application in healthcare.

5 CONCLUSION

This research has considerably explored the potential of blockchain generation in revolutionizing the data distribution mechanisms in healthcare. Through an intensive examination of the constraints inherent in conventional centralized structures, we have analyzed the transformative impact of blockchain's decentralized, transparent, and immutable nature. The proposed blockchain-primarily based method not only addresses the key worrying conditions of data integrity, security, and transparency but additionally introduces advanced interoperability and patient-centric data handling. Experimental results confirm the prevalence of the blockchain approach in terms of throughput, latency, scalability, and safety, indicating its readiness for actual-world implementation. This study reaffirms blockchain's ability in healthcare data control and also sets a basis for future exploration into its broader applications in various sectors. Future research may additionally focus on the practical deployment challenges, integration with present systems, and exploring the synergy among blockchain and different emerging technology which includes AI and IoT in healthcare.

REFERENCES

Abiteboul, S. and Stoyanovich, J. (2019). Transparency, fairness, data protection, neutrality: Data management challenges in the face of new regulation. *Journal of Data and Information Quality (JDIQ)*, 11(3):1–9.

Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., and Akin, E. (2023). A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions. *Electronics*, 12(6).

Attaran, M. (2022). Blockchain technology in healthcare: Challenges and opportunities. *International Journal of Healthcare Management*, 15(1):70–83.

Azaria, A., Ekblaw, A., Vieira, T., and Lippman, A. (2016a). Medrec: Using blockchain for medical data access and permission management. In *2016 2nd international conference on open and big data (OBD)*, pages 25–30. IEEE.

Azaria, A., Ekblaw, A., Vieira, T., and Lippman, A. (2016b). Medrec: Using blockchain for medical data access and permission management. In *2016 2nd international conference on open and big data (OBD)*, pages 25–30. IEEE.

Bertin, P., Bonjour, S., and Bonnin, J.-M. (2009). Distributed or centralized mobility? In *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*, pages 1–6. IEEE.

Bhutta, M. N. M., Khwaja, A. A., Nadeem, A., Ahmad, H. F., Khan, M. K., Hanif, M. A., Song, H., Alshamari, M., and Cao, Y. (2021). A survey on blockchain technology: Evolution, architecture and security. *Ieee Access*, 9:61048–61073.

Chen, Z., Xu, W., Wang, B., and Yu, H. (2021). A blockchain-based preserving and sharing system for medical data privacy. *Future Generation Computer Systems*, 124:338–350.

Chenthara, S., Ahmed, K., Wang, H., Whittaker, F., and Chen, Z. (2020). Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology. *Plos one*, 15(12):e0243043.

Chukmaitov, A., Harless, D. W., Bazzoli, G. J., Carretta, H. J., and Siangphoe, U. (2015). Delivery system characteristics and their association with quality and costs of care. *Health care management review*, 40(2):92–103.

Erkin, Z., Mennink, B., Nateghizad, M., and Maulany, C. Privacy-preserving distributed access control for medical data. In *Samarati, P.(ed.), Proceedings of the 15th International Joint Conference on e-Business and Telecommunications - SECRYPT, July 26-28, 2018, in Porto, Portugal*, pages 322–331, 2018. Setubal: SCITEPRESS.

Haleem, A., Javaid, M., Singh, R. P., Suman, R., and Rab, S. (2021). Blockchain technology applications in healthcare: An overview. *International Journal of Intelligent Networks*, 2:130–139.

Hildebrandt, M., Verheul, E., Jacobs, B., Meijer, C., and de Ruiter, J. (2016). Polymorphic encryption and pseudonymisation for personalised healthcare: A whitepaper.

Hu, M., Ren, Y., and Chen, C. (2023). Privacy-preserving medical data-sharing system with symmetric encryption based on blockchain. *Symmetry*, 15(5):1010.

Huang, H., Zhu, P., Xiao, F., Sun, X., and Huang, Q. (2020). A blockchain-based scheme for privacy-preserving and secure sharing of medical data. *Computers & Security*, 99:102010.

Hugoson, M.-Å. (2009). Centralized versus decentralized information systems: A historical flashback. In *History of Nordic Computing 2: Second IFIP WG 9.7 Conference, HiNC2, Turku, Finland, August 21-23, 2007, Revised Selected Papers 2*, pages 106–115. Springer.

Inclusion Cloud (2023). Centralized vs. decentralized data: Unveiling the great debate.

Kumar, R. and Bhatia, M. (2020). A systematic review of the security in cloud computing: data integrity, confidentiality and availability. In *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, pages 334–337. IEEE.

Laoutaris, N. (2018). Data transparency: Concerns and prospects [point of view]. *Proceedings of the IEEE*, 106(11):1867–1871.

Liang, X., Shetty, S., Tosh, D., Kamhoua, C., Kwiat, K., and Njilla, L. (2017). Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In *2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*, pages 468–477. IEEE.

McIntosh, D. (2018). We need to talk about data: how digital monopolies arise and why they have power and influence. *J. Tech. L. & Pol'y*, 23:185.

Mohammed, N., Fung, B. C., Hung, P. C., and Lee, C.-K. (2010). Centralized and distributed anonymization for high-dimensional healthcare data. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 4(4):1–33.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Decentralized business review*.

Newaz, A. I., Sikder, A. K., Rahman, M. A., and Uluagac, A. S. (2021). A survey on security and privacy issues in modern healthcare systems: Attacks and defenses. *ACM Transactions on Computing for Healthcare*, 2(3):1–44.

Pandey, A. K., Khan, A. I., Abushark, Y. B., Alam, M. M., Agrawal, A., Kumar, R., and Khan, R. A. (2020). Key issues in healthcare data integrity: Analysis and recommendations. *IEEE Access*, 8:40612–40628.

Rajput, A. S. and Balasubramanian, R. (2021). Privacy-preserving distribution and access control of personalized healthcare data. *IEEE Transactions on Industrial Informatics*, pages 1–1.

Salmon, J. and Myers, G. (2019). Blockchain and associated legal issues for emerging markets.

Vishwa, A. (2021). *MediBlock-A Privacy-aware Blockchain to store patients data and effective diagnosis methods*. PhD thesis.

Yue, X., Wang, H., Jin, D., Li, M., and Jiang, W. (2016). Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. *Journal of medical systems*, 40:1–8.

Zhang, Y. and Wen, J. (2017). The iot electric business model: Using blockchain technology for the internet of things. *Peer-to-Peer Networking and Applications*, 10:983–994.