International Journal of Web of Multidisciplinary Studies
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
website: https://ijwos.com OPEN ACCESS E-ISSN : 3049-2424
Volume 3 Issue 3 , March 2026
DOI: https://doi.org/10.71366/ijwos03032604156

MEDICHAIN : BLOCKCHAIN & MACHINE LEARNING HYBRID FRAMEWORK FOR SECURE MEDICAL RECORD SHARING

SURESHKUMAR T[1], BARKATH AYISHA A[2], JAYASHREE J[3], JEEVITHA S[4], THULASI S[5]
[1] *ASSOCIATE PROFESSOR, DEPARTMENT OF IT, MPNMJ ENGINEERING COLLEGE*
[2,3,4,5] *STUDENTS, DEPARTMENT OF IT, MPNMJ ENGINEERING COLLEGE.*

Article Info

Article History:

Published:30 March 2026

Publication Issue:
Volume 3, Issue 3
March-2026

Page Number:
555-560

Corresponding Author:
SURESHKUMAR T

Abstract:

The digitization of Electronic Medical Records (EMRs) has dramatically improved operational efficiency within healthcare institutions but has introduced significant vulnerabilities regarding data privacy, structural integrity, and unauthorized systemic access. Traditional centralized architectures present a singular point of failure, leaving sensitive records prone to silent tampering and large-scale data breaches. MediChain is a proposed multi-layer security paradigm intended to mediate these issues without the high overhead of public ledgers. By integrating a custom Python-based permissioned blockchain to ensure an immutable audit trail, coupling it with AES-256-CBC local encryption for robust off-chain data confidentiality, and deploying a real-time Isolation Forest machine learning model, MediChain synthesizes a comprehensive defensive ecosystem. The platform functions as a unified framework that enforces patient-governed smart-contract consent while mathematically isolating anomalous behavioral access patterns. This paper presents the architecture and efficacy of this hybrid methodology, demonstrating an effective, decentralized-adjacent solution designed specifically for restricted-access clinical networks.

Keywords: Electronic Medical Records, Permissioned Blockchain, Machine Learning, Isolation Forest, AES-256-CBC, Access Control, Smart Contracts, Cybersecurity

1. INTRODUCTION

Medical records contain some of the most private and sensitive information that a person generates throughout their lifetime. They contain personal identity details, diagnostic reports, and physical examination findings that demand the highest level of legal protection under frameworks such as the HIPAA, GDPR, and the Digital Personal Data Protection Act. Despite these strict obligations, healthcare ecosystems globally are confronted by a persistent challenge: traditional centralized database architectures remain acutely vulnerable to unauthorized access, insider threats, and systemic data tampering.

Current healthcare information systems operate as single points of failure. If an intruder or malicious database administrator gains valid credentials, they can alter clinical notes or download mass data without leaving a detectable cryptographic trace. Furthermore, standard single-factor authentication mechanisms fail to distinguish between a legitimate physician and an attacker utilizing stolen credentials, as these systems lack the concept of normal operational behavior.

The primary objective of this study is to present MediChain, a framework that merges cryptographic integrity with behavioral intelligence into a single cohesive platform. Rather than relying on public blockchains that incur

prohibitive costs and latency, MediChain utilizes a local permissioned blockchain to lock data integrity. It cross-correlates file access attempts with historical user behavior using an advanced unsupervised machine learning model to dynamically block unauthorized access.

2. LITERATURE REVIEW AND RELATED WORK

The integration of digital security mechanisms into clinical administration has evolved significantly. However, existing solutions present severe limitations, primarily revolving around the mutability of traditional databases and the computational bloat of modern decentralized ledger technologies.

Recent literature highlights the rapid adoption of blockchain technology for data immutability. Several modern implementations, such as MedRec (Azaria et al., 2016) [7], utilize Ethereum smart contracts to manage access permissions. However, the Ethereum Virtual Machine (EVM) scales poorly under the intense load of heavy medical records (e.g., DICOM files mapping MRI scans). Public nodes require transaction fees (gas) and compute every transaction line, creating extreme bottlenecks. Furthermore, placing health semantics on a public ledger introduces a fundamental privacy conflict.

In the realm of active cybersecurity, traditional supervised machine learning models fail because novel zero-day attacks and insider threats often lack labeled training data. Consequently, anomaly detection utilizing unsupervised algorithms like the Isolation Forest (Liu, Ting, and Zhou, 2008) [3] has emerged as an optimal pathway. By isolating anomalies rather than profiling normal data points, these algorithms mathematically identify deviations—such as abnormal login hours or erratic IP routings—flagging compromised accounts before data exfiltration occurs. The proposed MediChain framework addresses the fragmentation of these technologies by unifying off-chain storage, permissioned distributed ledgers, and intelligent behavioral tracking into one operational application.

3. PROPOSED SYSTEM ARCHITECTURE

The proposed system architecture is designed around a highly modular, multi-tier Flask paradigm. The architecture separates the concern of data confidentiality from data integrity and access-pattern security:

• **Presentation Tier:** A responsive web application built with Jinja2 templates, delivering role-specific dashboards (Admin, Doctor, Patient) alongside secure OTP-MFA authentication interfaces.

• **Application Logic Tier:** A Flask/Python backend handling RESTful requests, JSON Web Token (JWT) validation, and discrete functional blueprints. This layer includes three primary service components: the AESCipher service, the Blockchain service, and the AnomalyDetector service.

• **Data Tier:** A secure MySQL database managed through the SQLAlchemy Object-Relational Mapping (ORM) framework, mapping user records, blockchain blocks, access logs, and patient consent states.

• **Storage Tier:** A localized, off-chain file directory containing heavily encrypted (.enc) representations of the raw medical data, mathematically decoupled from the database state.

Data flows sequentially. When a physician requests a patient file, the Python backend triggers parallel asynchronous validations: verifying the active consent state in MySQL, recomputing the SHA-256 blockchain

linkages, and querying the in-memory Isolation Forest model. If all parameters align, the system streams the decrypted file to the authenticated client.

4. METHODOLOGY

4.1 User Management and Smart Contract Simulation

The methodology begins with strict identity governance. Passwords provided during registration are processed through the bcrypt hashing algorithm, which applies a one-way hash combined with a randomly generated salt. Post-login, multi-factor authentication (OTP) is engaged before issuing an HTTP-only JWT to mitigate Cross-Site Scripting (XSS).

Patient consent is governed dynamically without the latency of public blockchain virtual machines. A simulated smart contract (ConsentManager class) maps state-transition logic within the ORM. State changes reflect immediately, generating non-repudiable audit logs and hard-blocking requests from non-consented healthcare providers.

4.2 Off-Chain Storage and AES-256 Encryption

To prevent blockchain bloat, raw medical files are passed through the Advanced Encryption Standard (AES) operating in Cipher Block Chaining (CBC) mode with a 256-bit key structure. A unique 16-byte Initialization Vector (IV) ensures that identical plaintexts result in diverging ciphertexts:

$$C_i = E_K(P_i \oplus C_{i-1})$$

Where $C_0 = IV$, P_i represents the plaintext block segments, and E_K represents the AES encryption cipher. The raw file is never stored on disk.

4.3 Permissioned Blockchain and Proof of Work

Integrity is enforced by generating a SHA-256 hash of the original plaintext file and storing it as a metadata payload within a localized blockchain block (B_n). Each block is linked to its predecessor:

$$H(B_n) = \text{SHA256}(\text{Index}_n || \text{Timestamp}_n || \text{Data}_n || H(B_{n-1}) || \text{Nonce}_n)$$

To resist spam and validate network transactions, a Proof-of-Work (PoW) consensus mandates that the nonce iterates until the resulting hash achieves a predefined target difficulty (e.g., two leading zeros):

$$\text{PoW Valid} \iff H(B_n) < \text{Target\_Difficulty}$$

4.4 Machine Learning Threat Detection

The system extracts six behavioral feature vectors from the access_logs table: access frequency, unique IPs, average access hour, failed attempts, time variance, and weekend access flags. These vectors are fed into an Isolation Forest model to calculate an anomaly score:

$$\text{Anomaly Score}(x, n) = 2^{-\frac{E(h(x))}{c(n)}}$$

Where $h(x)$ is the path length to isolate instance x, and $c(n)$ is the average path length of unsuccessful search. Scores below -0.3 explicitly trigger a "Critical" state, immediately terminating the connection layer.

5. IMPLEMENTATION AND RESULTS

The proposed framework was implemented utilizing Python 3.9, establishing high throughput and immediate cryptographic confirmation.

5.1 Multi-Stage Access Pipeline Execution

The system was evaluated against a strict sequential four-phase pipeline triggered upon any file retrieval request:

1. Role and MFA Identity Verification.
2. Consent State Interrogation (Simulated Smart Contract).
3. Double-Sided Cryptographic Verification (SHA-256 cross-referencing).
4. Behavioral Clearance Matrix (Isolation Forest Prediction).

5.2 Threat Evaluation and System Resilience

The architecture was tested under simulated cyber-threat vectors. In instances of Internal Sabotage, where a database administrator covertly modified a clinical record, Phase 3 immediately flagged a mathematical discrepancy between the persistent blockchain metric and the calculated local file hash, dropping the transaction. In simulated Phishing attacks, where valid credentials were used asynchronously from foreign subnets, Phase 4 intercepted the request due to an extreme isolation score divergence, proving the efficacy of the predictive ML component.

5.3 Comparative Analysis

Testing indicates that the proposed digital mechanism significantly improves institutional security while maintaining near-instant latency compared to legacy paradigms:

Metric	Traditional EMR	Public Blockchain	MediChain Paradigm
Data Integrity	Dependent on Mutable DB Logs	Fully Immutable (On-Chain)	Fully Immutable (Permissioned Hash Chain)

Operational Speed	Near-Instant	High (Seconds to Minutes)	Near-Instant
Transaction Cost	None	Gas Fees (Variable)	None
Privacy	Transport Level	Pseudonymous	Full AES-256 Local Encapsulation
Anomaly Intelligence	Passive Logging Review	Rarely Implemented	Active Isolation Forest Monitoring

6. DISCUSSION

The implementation of MediChain introduces significant advantages for clinical datacenters. By intertwining permissioned blockchain mechanics with AES encryption and unsupervised machine learning, the systemic vulnerability surface area is drastically reduced. The framework successfully demonstrates that zero-trust security does not mandate the adoption of unwieldy public distributed ledgers or the expenditure of cryptocurrency transaction fees.

However, localized deployment introduces infrastructural dependency on internal server redundancy. Although the blockchain prevents modification, the physical deletion of localized encrypted nodes must still be mitigated by routine RAID backups. Furthermore, the Isolation Forest model requires periodic retraining on updated chronological data to prevent "concept drift" as hospital operational hours or IP ranges organically change.

Future research will focus on evolving the centralized Python blockchain implementation toward a broader Multi-Node Federated consensus design. Interlinking adjacent regional hospital domains utilizing strict P2P gRPC pathways will ensure global dataset validity extending far beyond an isolated single institution.

7. CONCLUSION

This study proposes a robust, multi-layer security paradigm for medical record management to address the trilemma of data tampering, unauthorized access, and lack of intelligent monitoring. MediChain successfully integrates an off-chain/on-chain permissioned architecture with dynamic smart-contract consent mapping and preemptive behavioral tracking. The proposed digital mechanism not only catalogs immutable access events but actively terminates anomalous connections in real time. By securing confidential datasets without sacrificing operational latency, MediChain provides a scalable, highly secure architecture for the future of decentralized electronic health ecosystems.

References

[1] Nakamoto, S. "Bitcoin: A peer-to-peer electronic cash system," 2008. Available: https://bitcoin.org/bitcoin.pdf

[2] Ekblaw, A., Azaria, A. et al. "A Case Study for Blockchain in Healthcare: MedRec," MIT Media Lab, 2016.

[3] Liu, F. T., Ting, K. M., and Zhou, Z.-H. "Isolation forest," Proc. IEEE 8th Int. Conf. Data Mining, pp. 413-422, 2008.

[4] National Institute of Standards and Technology. "Advanced Encryption Standard (AES)," FIPS PUB 197, Nov. 2001.

[5] Jones, M., Bradley, J., and Sakimura, N. "JSON Web Token (JWT)," IETF RFC 7519, May 2015.

[6] Breiman, L. "Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001.

[7] Azaria, A., Ekblaw, A., Vieira, T., and Lippman, A. "MedRec: Using Blockchain for Medical Data Access and Permission Management," 2nd International Conference on Open and Big Data (OBD), IEEE, 2016.

[8] R. S. S. Kumar et al., "Adversarial Machine Learning-Industry Perspectives," 2020 IEEE Security and Privacy Workshops (SPW), 2020.