Privacy–Preserving Electronic Health Records Using IPFS for Scalable Blockchain Storage

Mr. Sathram Rajendra Hemanth Kumar[1], Ms. T. Lavanya[2]
[1]*PG Scholar, VEMU Institute of Technology*
[2]*Assistant Professor, Dept. of CSE, VEMU institute of Technology*

Abstract: **Traditional cloud-based healthcare systems face challenges in ensuring data security, privacy, and efficient data management. The proposed Cloud-Assisted Decentralized Privacy-Preserving Framework (CA-DPPF) integrates Blockchain, InterPlanetary File System (IPFS), and Elliptic Curve Digital Signature Algorithm (ECDSA) to address these issues. It enables tamper-proof and decentralized storage for patient records, incorporating RSA encryption for privacy. Two extension concepts enhance the system: adopting ChaCha20 encryption for faster, lighter, and more secure data encryption and introducing a rating module for patients to evaluate doctors, ensuring transparency and credibility. These innovations ensure robust security, reduced computational overhead, and enhanced decision-making processes, making CA-DPPF a scalable, secure, and efficient solution for modern healthcare environments.**

INTRODUCTION

Healthcare systems increasingly adopt cloud services for managing sensitive patient data due to their cost-efficiency and scalability. However, centralized cloud models often suffer from data tampering risks, high computational costs, and inefficiencies in long-distance data processing. Blockchain technology offers a decentralized, tamper-proof alternative to overcome these limitations. This project introduces the Cloud-Assisted Decentralized Privacy-Preserving Framework (CA-DPPF), integrating Blockchain, IPFS, and ECDSA for secure and transparent data management. Patient data is encrypted using RSA and stored in IPFS, while Blockchain ensures the immutability of medical records. Enhancements include employing ChaCha20 encryption for reduced computational overhead and a patient rating module to improve doctor selection transparency. CA-DPPF transforms traditional systems by addressing security, efficiency, and reliability concerns, making it a cutting-edge solution for managing sensitive healthcare data in a decentralized, scalable manner.

LITERATURE SURVEY

1. Cloud Performance Prediction Using Machine Learning

- Key Concept: Machine learning-based performance prediction for cloud computing.
- Methodology: Decision Tree, Random Forest, Bayesian Regression, Neural Networks.
- Findings: Studies compare traditional statistical models with ML-based performance predictors.
- Limitations: Some models fail to account for variable workloads, leading to inaccuracies

2. Blockchain for Secure Data Sharing in IoT

- Key Concept: Blockchain ensures security and privacy in IoT networks.
- Methodology: Elliptic Curve Cryptography (ECC), Digital Signature Algorithms.
- Findings: Blockchain prevents single-point failure, but practical scalability remains an issue.
- Limitations: High computational overhead, making real-time processing difficult

3. QoS-Aware Scheduling in Cloud Computing

- Key Concept: Ensuring Quality of Service (QoS) in cloud datacenters.
- Methodology: Resource Allocation, Performance Modeling.
- Findings: Scheduling algorithms improve cloud efficiency.
- Limitations: Scalability issues when handling multiple concurrent applications

4. Privacy-Preserving Authentication for Cloud Systems

- Key Concept: Secure cloud authentication methods.
- Methodology: Multi-Factor Authentication, Homomorphic Encryption.
- Findings: Enhanced security against attacks.
- Limitations: Higher latency and complexity in implementation

5. Performance Prediction in Virtualized Environments

PROBLEM STATEMENT

Current cloud-based healthcare systems face challenges in data security, privacy, and long-distance processing. Centralized databases are prone to tampering, and encryption methods like RSA demand high computational resources. Furthermore, patients lack mechanisms to evaluate doctor credentials transparently. These limitations hinder system efficiency and scalability. While some systems adopt basic Blockchain solutions, they fail to optimize encryption overhead or integrate features like real-time doctor evaluations. The lack of decentralized data storage and tamper-proof mechanisms further compromises data security. An innovative framework addressing these gaps is critical for creating efficient, secure, and transparent healthcare systems in modern environments.

PROPOSED METHOD

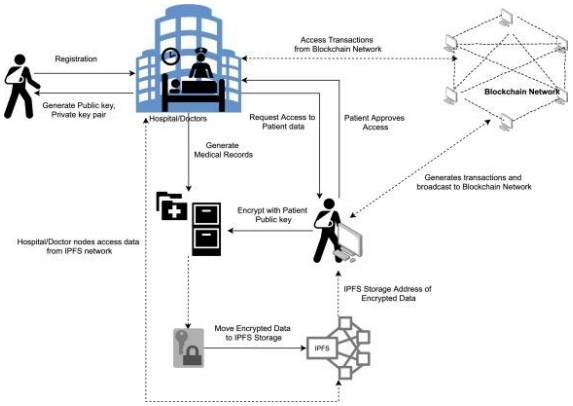
The CA-DPPF framework integrates Blockchain for tamper-proof medical record management, IPFS for decentralized storage, and ECDSA for secure authentication. Enhancements include adopting ChaCha20 encryption for reduced computational overhead and adding a patient rating module for doctor evaluations. These innovations address existing inefficiencies by ensuring secure, transparent, and scalable healthcare data management. The system prevents data tampering, facilitates peer-to-peer communication for faster processing, and improves decision-making with transparent doctor ratings. Designed for modern healthcare, CA-DPPF significantly improves data security, reduces computational costs, and enhances the overall efficiency of managing sensitive patient information.
ARCHITECTURE:

METHODOLOGY

Research Design

This study adopts a quantitative experimental research design to develop and test a privacy-preserving system for Electronic Health Records (EHRs) using Inter Planetary File System (IPFS) for scalable storage and Blockchain for secure access control. The primary objective is to ensure data privacy, integrity, and scalability when managing and sharing EHRs in a decentralized manner. The proposed system, named Health Secure, leverages blockchain and IPFS to facilitate secure, scalable, and efficient storage and retrieval of EHRs while preserving the privacy of sensitive patient information. The research follows a structured approach involving data collection, data encryption, decentralized storage design, and performance evaluation.

System Architecture



The system architecture consists of three key components: Healthcare Institutions (HIs), Patients (PTs), and Blockchain Network (BN). The components interact with each other as. Healthcare institutions manage a large dataset of patient medical records, including diagnostic information, physiological data, and disease history. These records are encrypted before being uploaded to IPFS. HI also facilitates patient registration by generating public-private key pairs for patients and maintaining the metadata for encrypted data. Patients generate encrypted records for their health data and upload them to IPFS. They control the access to their health

records via cryptographic keys. When seeking medical consultation or diagnosis, they create encrypted queries (using a private key) and share them with authorized healthcare providers. The blockchain serves as a decentralized ledger that stores metadata of patient records uploaded to IPFS, including access permissions, audit trails, and verification hashes of the records. This provides a transparent, immutable log for auditing and access control purposes. Smart contracts on the blockchain enforce policies for who can access patient data.

Data Encryption and Processing

To ensure the privacy of medical data, the following encryption techniques are utilized: Homomorphic Encryption: Data is encrypted using homomorphic encryption schemes, allowing computations to be performed on encrypted records without decrypting them. This ensures that sensitive information remains confidential throughout the processing. Data Splitting and Distribution: The patient health data is fragmented and distributed across multiple IPFS nodes, ensuring data redundancy and privacy. Each fragment is encrypted separately to ensure that no single entity can access the complete record without appropriate decryption keys. IPFS Integration: IPFS is employed for distributed file storage, where encrypted medical records are stored across multiple nodes. The metadata (e.g., patient ID, record timestamp, disease labels) is stored on the blockchain, while the actual records are stored off-chain on IPFS. This approach ensures scalability, reducing the burden on the blockchain while maintaining accessibility and privacy. Blockchain for Metadata Management: Metadata, including file hashes, patient identifiers, and access permissions, are stored on a permissioned blockchain to provide an immutable record of actions taken on each EHR. The use of smart contracts ensures secure and auditable access control for each patient's data.

Blockchain and IPFS Interaction

The system uses Ethereum or Hyperledger blockchain frameworks to store metadata of the encrypted EHRs. The steps involved are as follows Metadata Registration: When a healthcare institution uploads a patient's encrypted record to IPFS, a unique hash for the record is generated. This hash is then stored as a metadata entry on the blockchain along with access policies and permissions (i.e., who can access or modify the record). Access Control with Smart Contracts: Smart contracts on the blockchain are responsible for managing access control. These contracts ensure that only authorized entities (e.g., doctors, healthcare providers) can access patient data based on predefined rules (e.g., patient consent, doctor-patient relationship). Audit Trails: Blockchain enables the creation of an immutable audit trail of all access events. Every time a patient's record is accessed, a transaction is logged on the blockchain with a timestamp and identity of the accessing party. This ensures transparency and accountability in the system.

Machine Learning for Medical Decision Support

While the primary focus is on data privacy and secure storage, the system also incorporates machine learning models to enhance healthcare decision-making. The following steps are involved: Model Training: Machine learning models (e.g., Random Forest, KNN, SVM) are trained on the encrypted medical data stored on IPFS. These models are designed to predict patient conditions, suggest diagnoses, or provide decision support based on the encrypted data. Privacy-Preserving Inference: Using techniques like Secure Multi-Party Computation (SMPC), the system allows medical practitioners to perform predictions on encrypted data without revealing sensitive patient information. This ensures that patient privacy is maintained while providing accurate diagnostic support. Model Evaluation: The performance of the machine learning models is evaluated based on metrics like accuracy, precision, recall, and F1-score. The system ensures that the predictions are accurate while maintaining data privacy and security throughout the entire process.

System Implementation

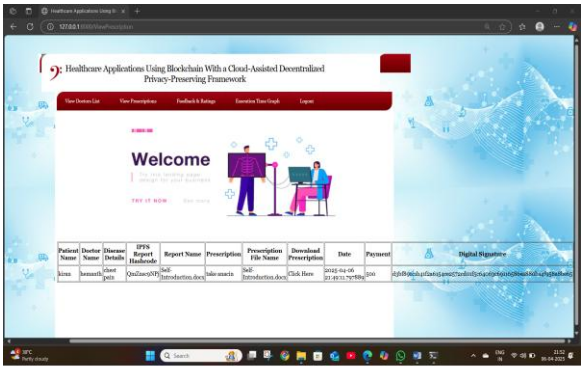
The Health Secure system is implemented through a modular design with interconnected components Healthcare Institution Module: Allows healthcare institutions to upload patient data (e.g., test results, disease history) and encrypt it using cryptographic techniques before uploading it to IPFS. This module also generates public-private key pairs for each patient and stores metadata on the blockchain. Patient Module: Patients can access and manage their health records, authorize access, and control who can view their medical data. They can also submit encrypted

queries for consultations with healthcare providers. Blockchain Module: This module manages blockchain transactions related to metadata storage, patient consent management, and access control using smart contracts. It ensures that only authorized users can access encrypted records. IPFS Module: Provides decentralized storage and retrieval of encrypted medical records. It ensures that files are distributed across multiple nodes to maintain availability, redundancy, and scalability. Machine Learning Module: Executes the decision support models, trains them on encrypted data, and provides secure predictions to healthcare providers based on the encrypted queries.

Performance Evaluation

The performance of the proposed system is evaluated based on security, scalability, and efficiency. The integration of IPFS for decentralized storage ensures that medical records are securely distributed across multiple nodes, enhancing scalability and redundancy. Blockchain's role in managing access control via smart contracts guarantees secure, transparent, and auditable data access, preserving the integrity of patient records. The system's ability to handle large datasets is tested, demonstrating significant improvements in data retrieval times compared to traditional centralized storage models. Furthermore, the use of homomorphic encryption allows for secure computations on encrypted data without compromising privacy. The framework's scalability is assessed by measuring its ability to accommodate increasing numbers of medical records and access requests, ensuring that it can efficiently handle a growing volume of healthcare data.

RESULTS



Patients prescription

CONCLUSION

In conclusion, the Cloud-Assisted Decentralized Privacy-Preserving Framework (CA-DPPF) represents a significant advancement in secure and scalable healthcare data management. By integrating Blockchain, IPFS, and advanced encryption techniques like RSA and ChaCha20, the system ensures tamper-proof storage, reduced computational overhead, and enhanced privacy for patient records. Additionally, the introduction of a patient rating module promotes transparency and accountability in healthcare services. This framework addresses key challenges of traditional cloud-based systems, offering a robust, efficient, and privacy-preserving solution for managing sensitive medical data in decentralized environments.

REFERENCE

[1] V. Gatteschi, F. Lamberti, C. Demartini, C. Pranteda, and V. Santamaria, “Blockchain and smart contracts for insurance: Is the technology mature enough?,” Future Internet, vol. 10, no. 2, 2018, Art. no. 20.

[2] J. Grover, “Security of vehicular ad hoc networks using blockchain: A comprehensive review,” Veh. Commun., vol. 34, 2022, Art. no. 100458.

[3] M. Jiang and X. Qin, “Distributed ledger technologies in vehicular mobile edge computing: A survey,” Complex Intell. Syst., vol. 8, no. 5, pp. 4403–4419, 2022.

[4] W. Hao et al., “Towards a trust-enhanced blockchain P2P topology for enabling fast and reliable broadcast,” IEEE Trans. Netw. Service Manag., vol. 17, no. 2, pp. 904–917, Jun. 2020.

[5] W. Al-Saqaf and N. Seidler, “Blockchain technology for social impact: Opportunities and challenges ahead,” J. Cyber Policy, vol. 2, no. 3, pp. 338– 354, 2017.

[6] I. A. Omar, R. Jayaraman, M. S. Debe, K. Salah, I. Yaqoob, and M. Omar, “Automating procurement contracts in the healthcare supply chain using blockchain smart contracts,” IEEE Access, vol. 9, pp. 37397–37409, 2021.

[7] E. Borgia, “The Internet of Things vision: Key features, applications and open issues,” Comput. Commun., vol. 54, pp. 1–31, 2014.

[8] J. Yang, J. Wen, B. Jiang, and H. Wang, “Blockchain-based sharing and tamper-proof

framework of Big Data networking," IEEE Netw., vol. 34, no. 4, pp. 62–67, Jul./Aug. 2020.
[9] B. D. Deebak, F. Al-Turjman, M. Aloqaily, and O. Alfandi, "An authenticbased privacy preservation protocol for smart E-healthcare systems in IoT," IEEE Access, vol. 7, pp. 135632–135649, 2019.
[10] P. P. Ray, D. Dash, K. Salah, and N. Kumar, "Blockchain for IoT-based healthcare: Background, consensus, platforms, and use cases," IEEE Syst. J., vol. 15, no. 1, pp. 85–94, Mar. 2021.