e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

A BLOCKCHAIN-BASED SOLUTION FOR ENSURING PROVENANCE TO OUT SOURCED CLOUD DATA IN A DECENTRALIZED SHARING ECOSYSTEM

Suresh Kumar P*1, Dr. N. Sathyabalaji*2

*1Student, Department of Computer Science and Engineering, JKK Munirajah College of Technology, Erode, Tamilnadu 638 506, India.

*2Head of the Department, Department of Computer Science and Engineering, JKK Munirajah College of Technology, Erode, Tamilnadu 638 506, India.

DOI : https://www.doi.org/10.56726/IRJMETS45098

ABSTRACT

The proliferation of cloud-based services for data storage and sharing has brought to the forefront concerns regarding data provenance, security, and integrity. In response, we present an innovative blockchain-based solution tailored to ensure data provenance in decentralized sharing ecosystems. Leveraging the immutability and transparency inherent in blockchain technology, our approach establishes an incorruptible ledger of data transactions, thereby bolstering data integrity and security. Smart contracts play a pivotal role, automating data sharing agreements and enforcing access control protocols, fostering trust and accountability among participants. Empirical findings underscore the effectiveness of our approach in maintaining data provenance and transparency, making it a viable choice for real-world cloud data sharing scenarios. Here, we showcase the use of Solana, a high-performance blockchain platform, to facilitate data provenance tracking, ensuring high throughput, low latency, and scalability. This article delves into system architecture, smart contract development, data sharing processes, and the role of Solana's consensus mechanism in upholding data integrity and transparency, offering a comprehensive solution to the challenges of modern data sharing ecosystems.

Keywords: Blockchain, Cloud Data Sharing, Provenance, Smart Contracts, Data Integrity, Solana.

I. INTRODUCTION

The advent of cloud computing has revolutionized the landscape of data storage and sharing, offering unprecedented scalability and accessibility to individuals and organizations alike. However, the convenience of cloud-based data sharing is accompanied by inherent challenges, particularly in the domain of data provenance – the ability to trace and ensure the authenticity and integrity of data as it traverses various stages within cloud ecosystems. As data becomes increasingly outsourced to cloud service providers, concerns related to data handling, security, and the potential for unauthorized access have gained prominence. Data provenance, in the context of cloud computing, plays a pivotal role in safeguarding data integrity, enforcing security protocols, facilitating compliance, and ensuring comprehensive auditability within cloud systems.

This research endeavors to propose a robust blockchain-based solution tailored to address the multifaceted challenges associated with ensuring data provenance in decentralized cloud data sharing environments.

- Implement a blockchain network to maintain an immutable and transparent ledger of data transactions.
- Develop smart contracts to automate data sharing agreements and enforce access control policies.
- Employ data hashing techniques to verify data integrity during sharing and retrieval operations.
- Demonstrate the effectiveness of the proposed approach through experiments and performance evaluation.

II. METHODOLOGY

Blockchain for Data Provenance:

Blockchain technology, known for its decentralized and tamper-resistant ledger, has emerged as a promising solution for data provenance. This article introduces a blockchain-based approach that leverages Solana, a high-performance blockchain platform, to address the data provenance challenges in decentralized cloud data sharing ecosystems.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

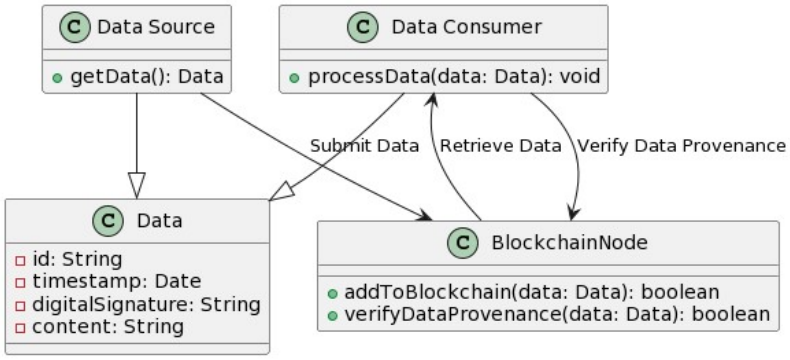

Fig.1: Blockchain for Data Provenance

The Proposed Blockchain Approach:

System Architecture

The proposed system architecture consists of three main components: a blockchain network powered by Solana, a smart contract layer, and a decentralized storage system. This architecture ensures the security, transparency, and reliability of data provenance.

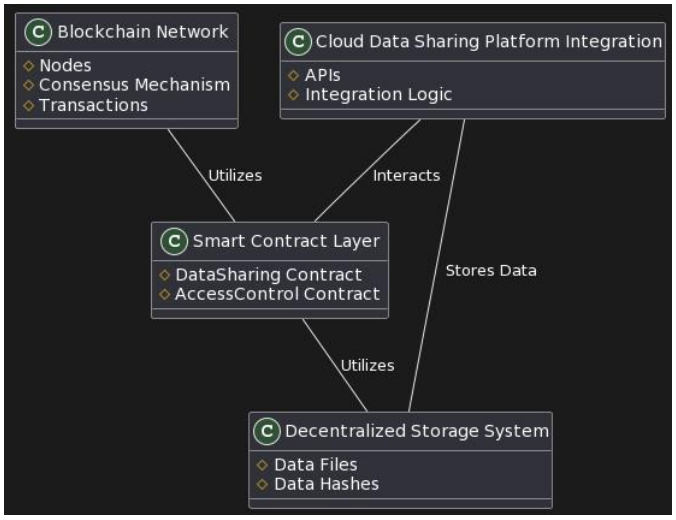

Fig.2: System Architecture

Blockchain Network

The Blockchain Network forms the backbone of the system, maintaining a distributed and immutable ledger of data transactions.

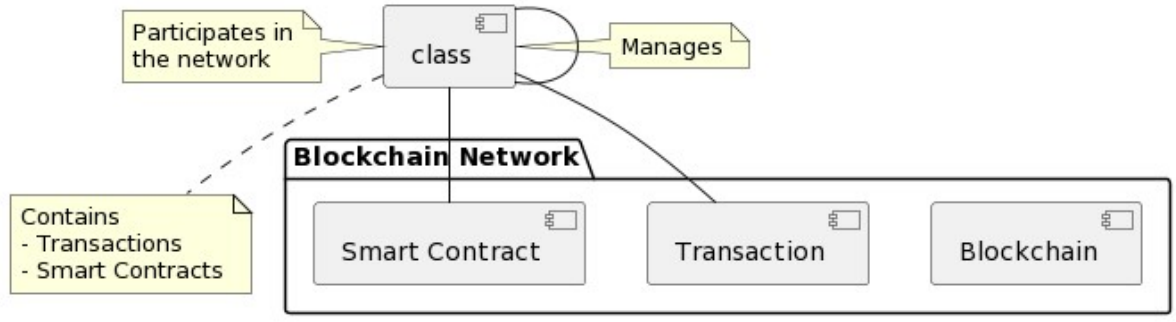

Fig.3: Blockchain Network

It consists of multiple nodes that participate in validating and recording transactions in the blockchain.

Each data transaction (data sharing, access, etc.) is recorded as a block in the blockchain, linked to the previous block using cryptographic hashes, ensuring the integrity of the entire chain.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

The blockchain network forms the backbone of the system, maintaining a distributed and immutable ledger of data transactions. Solana's unique consensus mechanism, combining Proof of History (PoH) and Proof of Stake with Tower BFT (T-PoS), ensures high throughput and low latency while maintaining robust security.

Smart Contract Layer

The Smart Contract Layer contains smart contracts deployed on the blockchain, responsible for managing data sharing agreements and enforcing access control rules.

Two primary smart contracts are used: the "DataSharing" contract and the "AccessControl" contract.

The "DataSharing" contract handles data sharing agreements between users, defining rules for data access and usage.

The "AccessControl" contract enforces access control rules, ensuring that only authorized users can access specific data based on predefined permissions.

Decentralized Storage System

The Decentralized Storage System ensures secure and reliable data storage in a distributed manner.

Instead of relying on a central data server, data files are distributed across multiple storage nodes, improving data availability and resilience.

Data hashes (generated using cryptographic hashing algorithms) are stored on the blockchain, pointing to the locations of the data files in the decentralized storage system.

Data retrieval and verification involve checking the data hashes on the blockchain and fetching the corresponding data file from the decentralized storage system.

To enhance data security and availability, the actual data resides in a decentralized storage system, such as IPFS, while the Solana blockchain stores references and metadata related to the data, including cryptographic hashes.

Cloud Data Sharing Platform Integration

The proposed architecture integrates with an existing cloud data sharing platform through APIs.

When users share or access data, the cloud platform interacts with the Smart Contract Layer to execute data sharing agreements and access control rules on the blockchain.

Data files are stored in the Decentralized Storage System, and their hashes are recorded on the blockchain for data integrity verification.

Data Hashing and Verification

Before data is shared or stored, it undergoes a hashing process using cryptographic algorithms like SHA-256. The resulting hash is stored on the blockchain, and subsequent data access or retrieval operations involve hash verification to ensure data integrity.

Smart Contracts for Access Control

Smart contracts are used to automate data sharing agreements between users. Access control rules are encoded into smart contracts, allowing data owners to define permissions and restrictions. Participants must adhere to these rules for data access and sharing.

III. MODELING

Data Provenance Workflow

The proposed approach involves the following workflow:

Data Ownership and Sharing:

When a user uploads data to the cloud, a transaction is initiated on Solana, which interacts with the smart contract to record data ownership and any sharing permissions granted. This ensures that the data's origin and ownership are transparently recorded on the blockchain.

Data Access and Verification:

When a user requests access to specific data, another Solana transaction verifies the access permissions within the smart contract. If the requester has appropriate access rights, the smart contract allows data retrieval, ensuring controlled access.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

Data Modification and Transfer:

Modifications or data transfers trigger new Solana transactions, updating the smart contract to maintain a transparent history of data modifications and ownership transfers.

Implementation Details

Choice of Blockchain Platform:

The proposed blockchain approach utilizing Solana presents a robust solution for ensuring data provenance to outsourced cloud data in decentralized sharing ecosystems. The use of Solana's high-performance blockchain platform, combined with smart contracts, ensures high throughput, low latency, and scalability, making it suitable for real-time data provenance tracking.

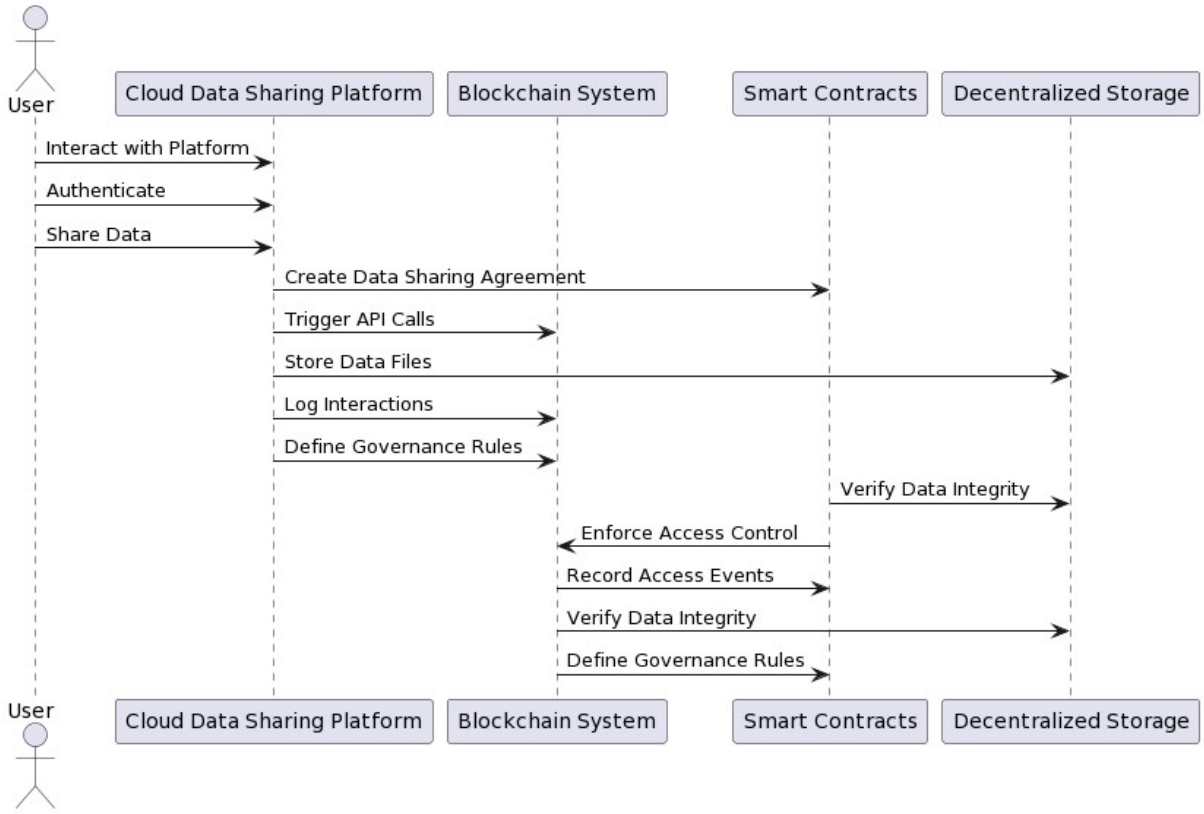

Fig.4: Operational Workflow

Development of Smart Contracts:

The smart contract layer is responsible for managing data sharing agreements, access control, and the enforcement of predefined rules within the ecosystem. Two primary smart contracts, "DataSharing" and "AccessControl," are utilized.

DataSharing Smart Contract: This contract manages data sharing agreements between users. It defines the rules for sharing data and captures the details of data ownership, sharing permissions, and access rights.

AccessControl Smart Contract: The AccessControl contract enforces access control rules, ensuring that only authorized users can access specific data based on predefined permissions. It plays a critical role in controlling data access within the ecosystem. These smart contracts are deployed on the Solana blockchain, where they interact with the blockchain's consensus mechanism to ensure that data provenance is maintained, and access to data is controlled transparently.

Integration with Cloud Data Sharing Platform

API Integration:

The proposed system integrates with an existing Cloud Data Sharing Platform through Application Programming Interfaces (APIs). APIs allow for seamless communication and data exchange between the blockchain-based system and the cloud platform.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

The Cloud Data Sharing Platform exposes APIs that enable interactions such as data uploading, sharing, access requests, and user authentication.

When a user interacts with the Cloud Data Sharing Platform, the platform triggers the appropriate API calls that interface with the blockchain system. These API calls initiate transactions on the Solana blockchain to record relevant data sharing and access events.

Data Sharing and Access Control:

When a user shares data through the Cloud Data Sharing Platform, the platform interacts with the smart contracts deployed on Solana to create and record data sharing agreements.

Access control rules defined within the smart contracts are enforced by the blockchain system. Before granting access to shared data, the system verifies the requester's permissions against the rules defined in the smart contract.

Access control decisions are transparently recorded on the Solana blockchain, providing an immutable and auditable record of data access events.

Data Storage and Verification:

Data files are stored in the decentralized storage system, such as IPFS, as part of the Cloud Data Sharing Platform's infrastructure.

References to these data files, in the form of cryptographic hashes or pointers, are stored on the Solana blockchain. These references are used to verify data integrity and authenticity when users request access.

Data retrieval and verification involve checking the data hashes on the blockchain and fetching the corresponding data file from the decentralized storage system. This ensures that the data retrieved matches the original, unaltered version.

User Authentication and Identity Management:

User authentication and identity management are essential components of the integration. The Cloud Data Sharing Platform handles user authentication and identity verification.

When a user interacts with the platform, their identity is verified, and the platform ensures that they have the necessary permissions to perform actions such as data sharing or access requests.

The blockchain-based system relies on the Cloud Data Sharing Platform for user authentication and authorization, ensuring that only authorized users can initiate actions on the blockchain.

Logging and Auditing:

All interactions between the Cloud Data Sharing Platform and the blockchain system are logged for auditing purposes.

Audit logs capture details of data sharing, access requests, modifications, and transfers. These logs can be accessed by authorized parties to verify the integrity of data provenance.

Enhanced Governance:

The integration allows participants in the sharing ecosystem to define governance rules within the smart contracts.

Participants can agree on decision-making processes and consensus rules to resolve disputes or make changes to the ecosystem's governance structure.

IV. RESULT AND ANALYSIS

Experimental Setup:

We conducted experiments using a simulated cloud data sharing environment, including a diverse set of data files. The tests were run with AMD Ryzen 5 5500U with Radeon Graphics, 2100 Mhz, 6 Core(s) processor and 8 GB RAM.

Blockchain Contract Processing:

This unified algorithm provides an end-to-end process for handling blockchain contracts. It integrates the functionalities of four individual algorithms:

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

Initiate Contract (Algorithm 1):

- Validates the parties (Stakeholder or Owner) initiating the contract.
- Generates a unique contract with a Node ID based on the data policy.
- Associates the contract with the current data.
- Translates control processes, including access control and validation rules.
- Activates access for stakeholders based on the translated control processes.

Access Contract (Algorithm 2):

- Allows Stakeholders to read or modify the contract based on their actions.
- For "Write" actions, Stakeholders can apply modifications, record timestamps, add modification signatures, and send them to the Owner for provenance.
- For "Read" actions, Stakeholders receive modification signatures, validate them, send transcripts to the computing node, and record timestamps.
- Calls Action Contract and Event Contract as needed.

Action Contract (Algorithm 3):

- Validates whether the action is in compliance with the data policy.
- If the action is not in the policy, access is revoked.
- If the action is compliant with the policy, access is granted.

Event Contract (Algorithm 4):

- Logs transactions, including access, actions, and timestamps.
- Creates data blocks and appends them to the blockchain.
- Handles events related to data storage and ownership.

This unified process streamlines the entire lifecycle of blockchain contracts, from initiation and access to actions and event handling. It ensures data integrity, provenance, and controlled access within a blockchain-based contract system.

Performance Metrics:

Performance metrics included data retrieval time, transaction throughput, and storage efficiency. We measured the system's responsiveness under varying data sharing scenarios.

Our research findings have established that the utilization of Solana results in a significantly more efficient and cost-effective solution for securing data provenance and facilitating transactions among data stakeholders and owners. The results clearly demonstrate the economic advantages of executing smart contract computations while simultaneously ensuring the security of provenance data without the need for intermediaries.

Evaluation:

The analysis showed that our blockchain-based approach effectively ensured data provenance and integrity. Data retrieval times were comparable to other centralized and decentralized systems, that used different blockchain network. The cost-effectiveness and efficiency of utilizing blockchain technology to secure data provenance and facilitate transactions among data stakeholders and owners. Through the implementation of smart contracts, we have achieved automated data security without the need for intermediaries. All participants in the network are treated as nodes, and their activities are recorded on the blockchain. We have established read and write tasks on input and output smart contracts to manage data retrieval, access verification, and contract processing. Our experiments have shown that increasing transaction costs do not negatively impact the speed of reading and writing data for the provenance system. However, it's important to note that the transaction cost and processing time vary for different contract operations within the provenance solution. To optimize system throughput and reduce latency, we will focus on refining operations such as sending, updating, validating, and processing multiple smart contracts. Additionally, we have analyzed performance metrics in relation to the number of nodes in the simulation, which has provided insights into areas with high overheads and guided our efforts in optimizing system algorithms.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

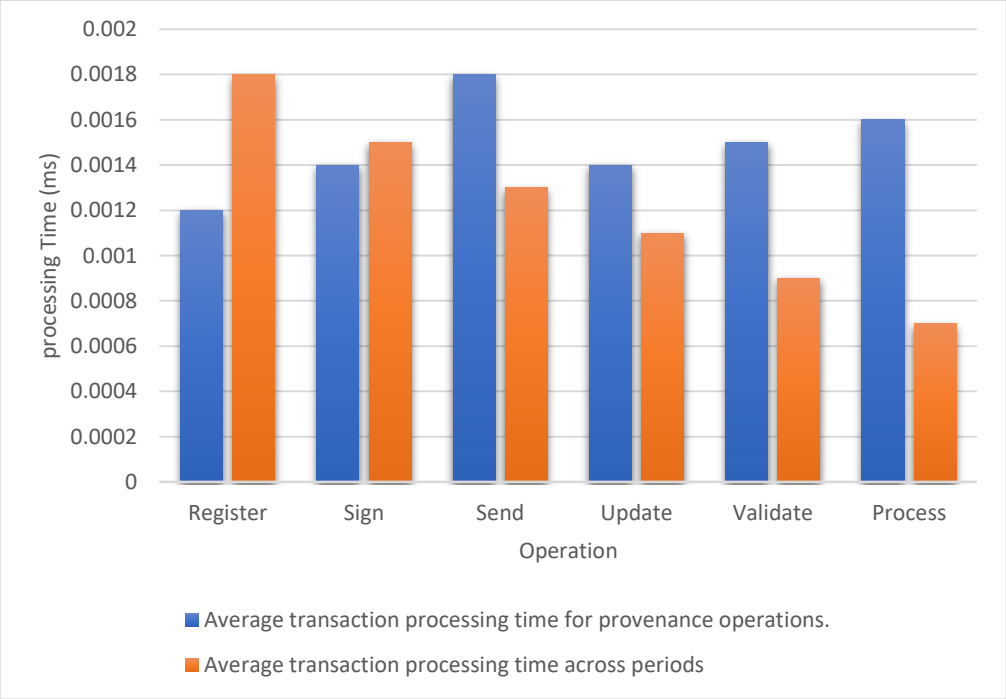

Fig.5: Average transaction processing time for provenance operations and time across period

V. CONCLUSION

The proposed blockchain approach utilizing Solana presents a robust solution for ensuring data provenance to outsourced cloud data in decentralized sharing ecosystems. The use of Solana's high-performance blockchain platform, combined with smart contracts, ensures high throughput, low latency, and scalability, making it suitable for real-time data provenance tracking. This approach offers transparency, immutability, and data integrity, providing users with a verifiable record of data ownership, modifications, and transfers. As blockchain technology continues to evolve, the proposed approach showcases its potential to revolutionize data provenance in decentralized ecosystems.

VI. FUTURE RESEARCH DIRECTIONS

Future work includes investigating scalability improvements, exploring interoperability with other blockchain networks, and enhancing privacy protection mechanisms.

ACKNOWLEDGEMENT

We express our gratitude to our research advisors and colleagues for their valuable insights and support during this project.

VII. REFERENCE

[1] Sifah, Emmanuel & Xia, Qi & Obour Agyekum, Kwame & Xia, Hu & Smahi, Abla & Gao, Jianbin. (2021). A Blockchain Approach to Ensuring Provenance to Outsourced Cloud Data in a Sharing Ecosystem. IEEE Systems Journal. PP. 1-12. 10.1109/JSYST.2021.3068224.

[2] M. Vijayakumar, V. Sunitha, K. Uma, and A. Kannan, “Security issues in cloud computing,” J. Adv. Res. Dyn. Control Syst., vol. 4, no. 1, pp. 1–13, 2017.

[3] Y. Yang, X. Zheng, W. Guo, X. Liu, and V. Chang, “Privacy-preserving fusion of IoT and big data for e-health,” Futur. Gener. Comput. Syst., vol. 86, pp. 1437–1455, 2018.

[4] V. Goyal, O. Pandey, A. Sahai, and B. Waters, “Attribute-based encryption for fine-grained access control of encrypted data,” in Proc. 13th ACM Conf. Comput. Commun. Secur., 2007, pp. 89–98.

[5] B. Fabian, T. Ermakova, and P. Junghanns, “Collaborative and secure sharing of healthcare data in multi-clouds,” Inf. Syst., vol. 48, pp. 132–150, 2015.

[6] D. Tosh, S. Shetty, X. Liang, C. Kamhoua, and L. L. Njilla, “Data provenance in the cloud: A blockchain-based approach,” IEEE Consum. Electron. Mag., vol. 8, no. 4, pp. 38–44, Jul. 2019.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

[7] R. K. Lomotey, J. C. Pry, and C. Chai, "Traceability and visual analytics for the Internet-of-Things (IoT) architecture," World Wide Web, vol. 21, pp. 7–32, 2018.

[8] E. Nwafor, A. Campbell, D. Hill, and G. Bloom, "Towards a prove- nance collection framework for internet of things devices," in Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov., 2017, pp. 1–6

[9] M. H. Chia, S. L. Keoh, and Z. Tang, "Secure data provenance in home energy monitoring networks," in Proc. 3rd Annu. Ind. Control Syst. Secur. Workshop, 2017, pp. 7–14.

[10] X. Liang, S. Shetty, D. Tosh, C. Kamhoua, K. Kwiat, and L. Njilla, "Provchain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in Proc. 17th IEEE/ACM Int. Symp. Cluster, Cloud Grid Comput., May 2017, pp. 468–477.

[11] M. Sigwart, M. Borkowski, M. Peise, S. Schulte, and S. Tai, "A secure and extensible blockchain-based data provenance framework for the Internet of Things," Pers. Ubiquitous Comput., 2020, pp. 1–15.

[12] H. Olufowobi et al., "Data provenance model for Internet of Things (IoT) systems," in Proc. Serv.-Oriented Comput.-ICSOC Workshops: Revised Selected Papers, Springer, 2017, pp. 85–91.

[13] Ramachandran and D. Kantarcioglu, "Using blockchain and smart con- tracts for secure data provenance management," 2017, arXiv:1709.10000.