Proof-of-Work for Merkle based Access Tree in Patient Centric Data

B Ravinder Reddy[1], T Adilakshmi[2]
Research Scholar, Department of CSE, UCE (A), OU[1]
Assistant Professor, Department of CSE, Anurag University, Hyderabad, T.S. India[1]
Professor and Head, Department of CSE, Vasavi College of Engg. (A), Hyderabad. T.S. India[2]

***Abstract*—With the advent of wearable devices and smart health care, the wearable health care technology for obtaining Patient Centric Data (PCD) has gained popularity in recent years. To establish access control over encrypted data in health records, Ciphertext Policy-Attribute Based Encryption (CP-ABE), is used. The most critical element is granting secure access to the generated information. However, with growing complexity of access policy, computational overhead of encryption and decryption process also increases. As a result, ensuring data access control as well as efficiency in PCD collected by wearables is crucial and challenging. This paper proposes and demonstrates a proof-of-work for the Merkle-based access tree using notion of hiding the sensitive access policy attributes.**

***Keywords*—Merkle tree; hashing; CP-ABE; access policy; PCD**

I. INTRODUCTION

In order to manage access to important or valuable resources, access control is governed in computer security, according to Di Francesco Maesa et al. [1]. When access requests are made, access control policies are matched to the existing access context to determine subjects' permissions to resources. Many access control methods have been established in research to prevent unauthorized access to vital assets that are more centralized than distributed, such as healthcare records. According to Hong-Ning Dai et al. [2], Healthcare frameworks will inevitably undergo a digital change. The Internet of Medical Things (IoMT) is critical, but inherent security and privacy concerns limit mainstream use. According to Vincent Hu et al., [3], Blockchain has provided new insights into the security of clinical information, with the highlights of distributed, trustless data storage, peer-to-peer transmission, and encryption algorithms, and it may overcome the inconsistency between information sharing and privacy with suitable security measures. Using EHR raises issues regarding the confidentiality, protection, and dependability of medical information, according to Liang Huang et al. [4]. According to Buket Yüksel et al., [5] the combination of Blockchain with cloud has drawn a lot of attention recently for ensuring effectiveness, transparency, security, and providing better cloud administrations. To maximize the capability of the Blockchain-cloud mix, a full understanding of the present initiatives in this field is essential. Since healthcare applications are a good example of dispersed environments and depend on servers for storage and processing, CP-ABE is one of the most promising security mechanisms in this area. One-to-many communication [6] is supported by CP-ABE, and the owner defines the access policy, such that only approved decryptor can access the data. An encrypted storage system is a use case for CP-ABE. One encryption key cannot encrypt many sets of data, preventing the implementation of fine-grained access control. The CP-ABE approach may provide access control without increasing the amount of keys by indicating the set of properties of the decryptor, such as affiliation. Enhanced schemes of ABE include broadcast encryption, multi-authority provisioning, and anonymous user or decryptor. Additionally, the number of attributes and pairing computations are inversely related. On the other hand, Merkle trees are designed to authenticate messages with a distinctive signature while also enabling an intended validator to confirm the authenticity of one message without exposing the validity of others.

A. Research Problem

One of the main challenges in managing access to sensitive patient information is the attribute disclosure problem. Traditional access control methods, such as role-based access control (RBAC) and access control lists (ACLs), are not well-suited to the attribute disclosure problem, as they do not provide fine-grained control over access to sensitive patient information based on the user's attributes. CPABE is one of the few models that can provide this type of access control, but the attribute disclosure problem still persists as the access policy is exposed. To overcome this issue, a novel approach of attribute hiding through Merkle tree is proposed for PCD while preserving privacy. The Proof-of-work ensures that the decryption is performed by an authorized user and not an attacker.

B. Research Questions

1) What are the main challenges in managing access to sensitive patient information?

2) How does the proposed Merkle tree based access structure for sensitive attributes of PCD in CPABE avoid unauthorized access and preserve privacy?

3) How can cryptographic techniques be used to ensure that the Merkle tree-based CPABE system is verifiable?

4) What are the benefits of a verifiable Merkle tree-based CPABE system in terms of in security and privacy?

C. Research Objectives

The research objective of this study is to address the main challenges in managing access to sensitive patient information by proposing a Merkle tree based access structure for sensitive attributes of PCD in CPABE, and examining how this

structure can be used to avoid unauthorized access and preserve privacy. Additionally, the objective is to investigate the use of cryptographic techniques to ensure that the Merkle tree-based CPABE system is verifiable and to explore the benefits of such a system in terms of security and privacy.

D. Research Significance

The Proof-of-work proposed in the article provides a secure method for managing sensitive attributes in a patient-centric approach. It enhances security and integrity for both users and providers in healthcare associations. The proposed scheme uses a monotonic access structure for policy transmission, and is resistant to chosen-ciphertext attacks (CCA) even when an attacker acquires multiple private keys. Additionally, it can handle large volumes of data and provide quick integrity proof. This model contributes new knowledge in the domain of secure access control for sensitive patient information.

II. Literature Review

By 2025, the worldwide IoT healthcare industry is anticipated to grow to 534.3 billion US Dollars. Carminati B, et al., [7] wearable health care is a popular and convenient technology that enables users to access medical services. Wearable medical data needs rapid and accurate information sharing from any place for better healthcare choices. However, the security and confidentiality of such patient information will become a major issue while sharing. When unauthorized access to information is unavoidable in medical settings, the security of patients' data is crucial. Siti Dhalila et al., [8] to overcome these challenges, new systems must be created with protected electronic health data, efficient storage, and a properly verified retrieval mechanism. Clauson et al., [9] aside from the patient-provider interaction and access-sharing, such a provision must also protect the patient's privacy with a greater focus on patient-centric health data management. Shan Jiang et.al., [10] presented an access control mechanism based on smart contracts for access verification to get EHRs sharing by integrating data dumping and sharing processes for providing e-medical services through cloud and Blockchain. The medical information obtained by IoT devices is forwarded to nearest edge workers for information management while providing security. The data is then shared via Blockchain exchange.

According to C. Nguyen et al., [11], a framework-proposed fast CP-ABE model can assign costly computing jobs to semi-trusted third parties while keeping consistent number of simple calculations. To validate the validity of the decryption output, a Boneh-Lynn-Shacham signature model is utilized. Wang et al., [12] the approach creates a privacy schema for clinical data that is based on Blockchain and the cloud. It can perform safe insurance and clinical information integrity verification, as well as handle computation, information sharing, and security challenges. Wang et al., [13] offer a patient-driven PCD sharing system to preserve patients' privacy and provide them control over their PCDs. Prior to reevaluating, all PCDs in this structure are encrypted with multi-authority and ABE, addressing the key facilitation issue and achieving fine-grained access control. Additionally, anonymous authentication between the cloud and the user is advised to preserve integrity while concealing the patient during validation. Leyou Zhang et al., [14] demonstrate a successful character identity-based distributed decryption technique for a medical records sharing framework. It allows them to share their data with different people without having to recreate the decryption of their private key. ABE was proposed [15] in order for the user to decode; there must be at least n attributes that match between the ciphertext and the user. V Goyal et al., [16] design extended the expressiveness of access structure by tying a ciphertext to a set of attributes, which they dubbed KP-ABE. Bethencourt et al., [17] introduced CP-ABE, which employs attributes to reflect the credentials of a user. Healthcare data that is sent in real-time or kept on a third-party cloud server is susceptible to a number of threats. A patient's life might be severely affected by the unauthorized access to sensitive medical information. Communication between multiple devices is one of the criteria for an IoT healthcare system. CP-ABE is a method that has promise for providing role-based access control in this situation. A type of public key cryptosystem called CP-ABE uses a collection of attributes to establish the user roles. In this case, the decryption key is linked to the receiver's attribute set and the access policy is contained with ciphertext. If and only if, the recipient's attribute set complies with the access structure/policy used for encryption, a successful decryption will take place. The data owners established the access policy. According to Hui Cui et al., [18] the decryption keys were only a collection of attributes with no tree structure. In a CP-ABE system, an access structure is included in the ciphertext, which may leak sensitive information about the underlying plaintext and the privileged receivers since anybody who views the ciphertext may learn the privileged recipients' attributes from the associated access structure.

Patients can encrypt data in a variety of methods to tighten access control in healthcare, such that anybody who reads the encrypted text can only comprehend the patient's publicly known qualities and the sensitive information stays hidden. This is according to Vijayan, V et al., [19]. Use of CP-ABE to encrypt the whole access policy with attributes or only the portion of the policy that has to be concealed is one or more solutions. However, the system cannot determine if the end user has the necessary authorization to access it because the ciphertext can only be retrieved by authorized end users, according to Nishide, T et al., [20]. This method's inability to outsource the decryption costs is its second drawback. The system will not be protected by CP-ABE encryption if the policies are wholly or partially obscured. Data must be patient-centric in a healthcare setting since it may be shared across several domains. This may be accomplished by hiding the access policy property using CP-ABE's different degrees of hashing, which supports fine-grained access control. It must also be done efficiently and securely to confirm the existence of the attributes used in the formulation of the access policy without disclosing the content at the time of decryption.

III. Preliminaries

The access tree is a list of all stakeholders who have access. The proposed approach's major purpose is toprove the existence of attributes in the access structure without revealing the sensitive attributes to the decryptor.

A. System Model

CP-ABE, according to Wang, Shulan, et al. [12], can enable privacy preserving and safe data sharing in public environment. The policy is embedded in the ciphertext, and each user's private key is based on attributes set and successful decryption can occur, iff the ciphertext or attributes satisfy the access key. Fig. 1 shows the Architecture for CP-ABE policy.

B. Access Tree for PCD

The CP-ABE access policy with attributes like hospital, physician, Insurance company and department is defined for accessing PCD as in figure 2 is ("Physician" AND "Hospital X") OR ("Insurance company X" AND "Insurance Department"). The proposed algorithm has five phases:

Phase 1: Initial Setup phase - $S_e(\lambda, S) \rightarrow (PU_K, MS_K)$

The K_{GC} generates PU_K and MS_K, respectively, given by a λ and S.

Phase 2: Key Generation phase - $K_G(PU_K, MS_K, S) \rightarrow S_K$

For M. PU_K, MS_K and S, the K_{GC} generates S_K.

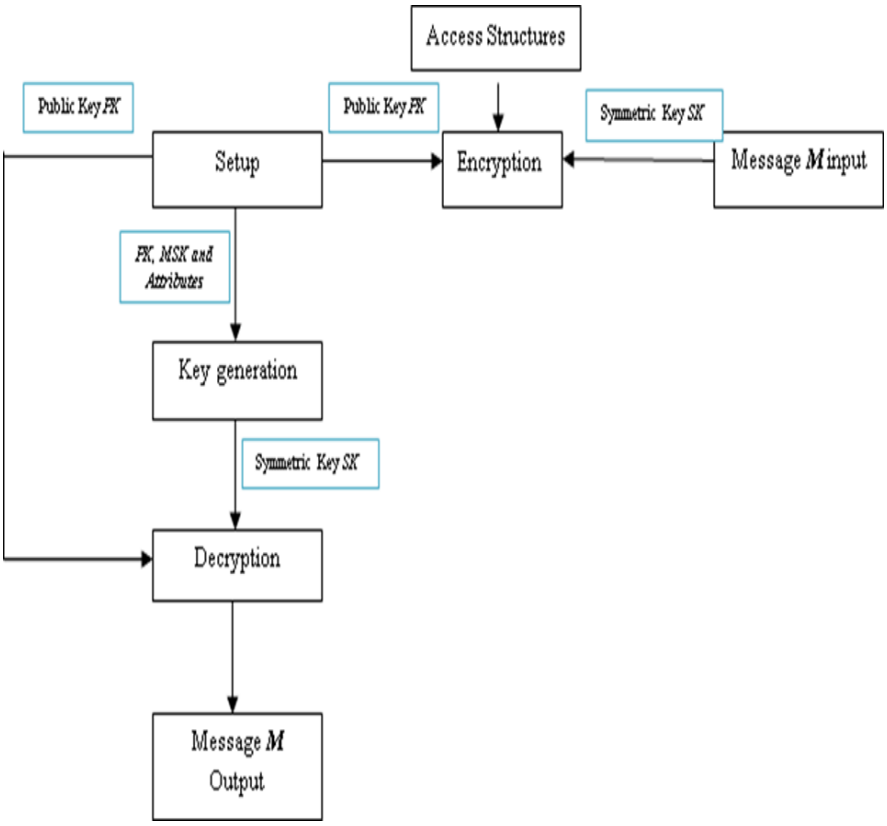

Fig. 1. Traditional CP-ABE architecture.

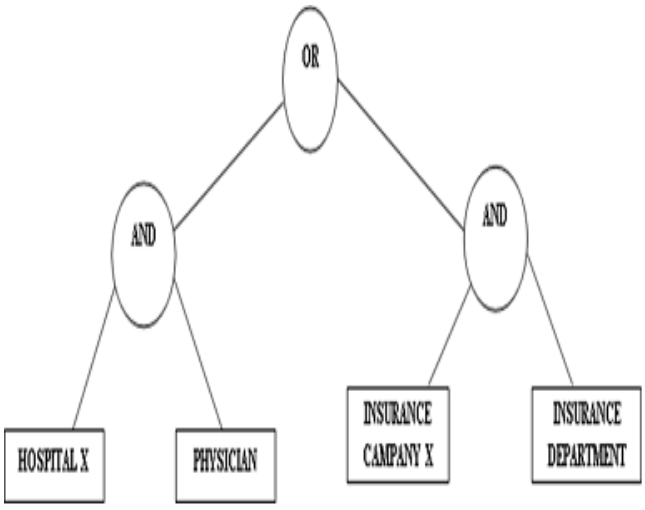

Fig. 2. The access tree generated for PCD.

Phase 3: Access Tree Construction phase

Given a collection of messages, M = {m1, m2,..., mn} creates a Merkle tree whose interior nodes include the concatenation of the hash values corresponding to its leaf nodes and whose leaves include the hash value of each message m in M. In the Fig. 3 below, shows the Merkle tree construction through attributes set, A= {a_1, a_2, a_3, a_4}. H_R represents the root hash value also called Merkle Root.

Phase 4: Encryption phase - $E(PU_K, S_T, NS_T, M) \rightarrow C_T$

The sender outputs the CT with input PUK, trees ST, NST, and M, where ST is a sensitive Merkle tree and

NST is a non-sensitive public tree.

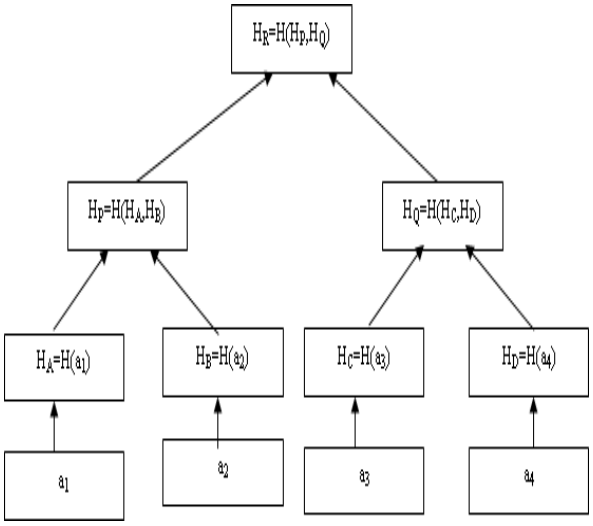

Fig. 3. The sensitive attribute Merkle tree S_T.

Phase 5: Decryption phase - $D(C_T, S_K) \rightarrow M$

The decryptor then decrypts C_T with S_K and outputs message M. By moving up the tree T_S from a specific transaction node to root node, Merkle proofs can demonstrate that a specific attribute is present during decryption. The decryptor examines the public tree T_{NS} and discovers the Merkle proof for S_T to verify the integrity of hidden attributes.

Consider a_3 from the attribute set A and $H`_C$ the hash of $a_3$ to be verified, represented by bolded rectangle. Given $H`_P$ from other sub tree and $H`_D$, the adjacent node hash value, represented by dashed rectangles. The $H`_R$ value can be calculated. If $H`_R == H_R$, the hash root value of the tree when generated as in Fig. 4, the attributes are considered to be secure.

C. Data Set

The tree was designed for use in experiments using the IBM-Clinical Hub [21] data set, which was stated in [23]**.** This data set contains clinical data elements that, when paired with patient identification data, allow for the storage, production, and access of longitudinal patient records. The IBM-Clinical Hub is a more comprehensive workbench paradigm that offers access to consuming systems and processes as well as patient and clinical data. An ongoing stream of HL7 events and messages is used to construct the longitudinal patient record.

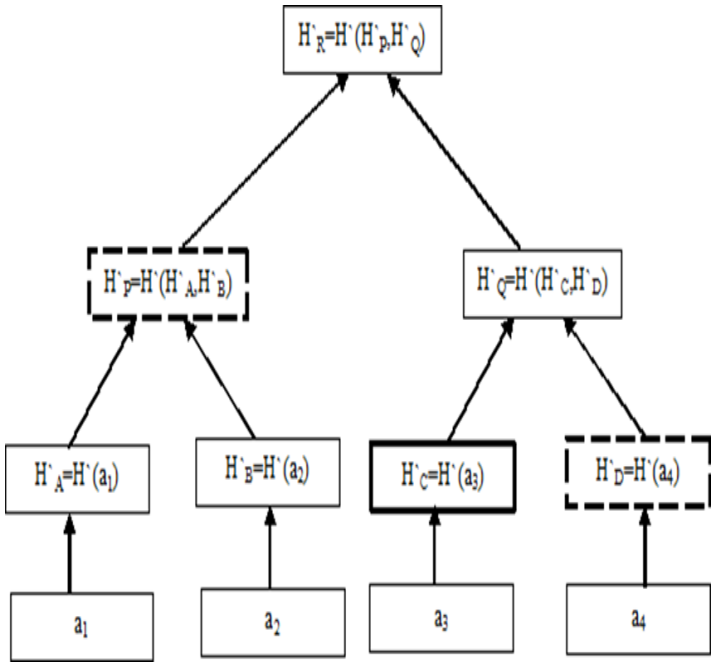

Fig. 4. The proposed sensitive attributes Merkle tree with verifiable nodes.

IV. Proposed Verifiable Merkle Proof Model for Sensitive Attributes

Merkle proof is a type of selective disclosure and is a useful strategy for protecting users' privacy by hiding unwanted sections of certificates or other documents and only offering partial disclosure for evidence. A Merkle proof validates specific transactions represented by a Merkle hash root's leaf or branch hash. Table II represents the list of PCD sensitive attributes for which the tree is generated with the root, r. The decryptor in the proposed Merkle tree based CP-ABE sensitive attribute access structure scheme will verify the integrity of the non-public attributes by selecting one attribute at random and by giving its hash value as input for the proof. In turn the decryptor will be given few other branch nodes hash values at random. Based on these multiple available hash values, the decryptor constructs the tree resulting to root hash. If the final hash value matches Merkle tree root, it means the attribute exists and the transaction can be processed.

Algorithm for Proposed Model

Step 1:
Consider the access tree [20] ST with root R. Hence, ST == TR.
Where SsT represents the subtree of sensitive attributes tree ST rooted at the node ST.

Step 2:
If A(β) ∈ SsT, Then denote [15] it as SsT(β) = 1.

Step 3:
Compute SsT(β) recursively as follows.
if(ST ≠ leaf)
Evaluate SsT′(β) = 1 ∀ children ST′ of node ST.
SsT(β) ==1
iff ≥ ks with children 1.
if(ST == leaf),
Then SsT(β) 1
iff A(ST) ∈ β.

V. Results and Discussion

The proposed Merkle tree based sensitive attribute access structure in CPABE is implemented using the pymerkle package[22], and its performance is evaluated by measuring the time taken to generate the Merkle tree as in Table I and Merkle proof for varying numbers of sensitive attributes using the SHA-256 algorithm. As the number of sensitive attributes increases, the size of the tree and its height also increases, resulting in longer generation times. Additionally, the results show that the Merkle proof computational time increases with the number of sensitive attributes. However, the time taken to generate the tree and Merkle proof decreases as the key size for the SHA-256 algorithm increases. These results demonstrate that the proposed scheme as in Figure 5 is efficient in terms of computational time and can be effectively implemented in real-world scenarios.

A. Comparison with other Frameworks

The Table II represents the comparison of the proposed model with other schemes, covering CPU load, data integrity, proof-of-work, efficiency and privacy. By creating a Merkle tree proof, guaranteeing integrity via a hash and providing effective data access control, the proposed CP-ABE scheme is capable to implement hidden sensitive attribute access policies with less computation load

B. Security Analysis

1) Chosen-ciphertext attack: The proposed Merkle tree model is secure against chosen-ciphertext attack is through the collision-resistant scheme. This means that it is computationally infeasible for an attacker to find two different sets of data blocks that have the same Merkle root.

Let H(x) be a cryptographic hash function that takes in input x and produces a fixed-length output. We can construct a Merkle tree with n leaf nodes, where each leaf node i is a hash of a data block Di, represented as H(Di). The non-leaf nodes in the tree are the hash of their children nodes, represented as H(Li, Ri) where Li and Ri are the left and right child nodes respectively.

The Merkle root of the tree, represented as R, is the hash of the root node and it is computed as R = H(H(L1, R1), H(L2, R2) … H(Ln/2, Rn/2)). To show that the tree is collision-resistant, assume that there exists an attacker who can find two different sets of data blocks, D1 and D2, such that H(D1) = H(D2) for all i. Therefore, the attacker can construct two different Merkle trees, T1 and T2, with the same Merkle root.

However, since H(x) is a collision-resistant hash function, it is computationally infeasible for the attacker to find two different inputs x1 and x2 such that H(x1) = H(x2). Therefore, it is also computationally infeasible for the attacker to find two different sets of data blocks that have the same Merkle root, and the Merkle tree structure provides a secure method for verifying the integrity of the sensitive attributes in the access structure. Additionally, the use of a one-way hash function in the construction of the Merkle tree ensures that it is computationally infeasible for an attacker to obtain the original sensitive attributes from the hash values stored in the tree, providing further security against chosen-ciphertext attacks.

TABLE I. Tree Generation Time Vs Proof Computational Time

S.No.	Number of Sensitive Attributes	Size/ Total nodes	Tree Height	Merkle Tree generation time (in seconds)*10^{-2}	Merkle Proof computational time (in seconds) *10^{-2}
1.	24	47	5	97.6	23.4
2.	90	179	7	33.0	30.1

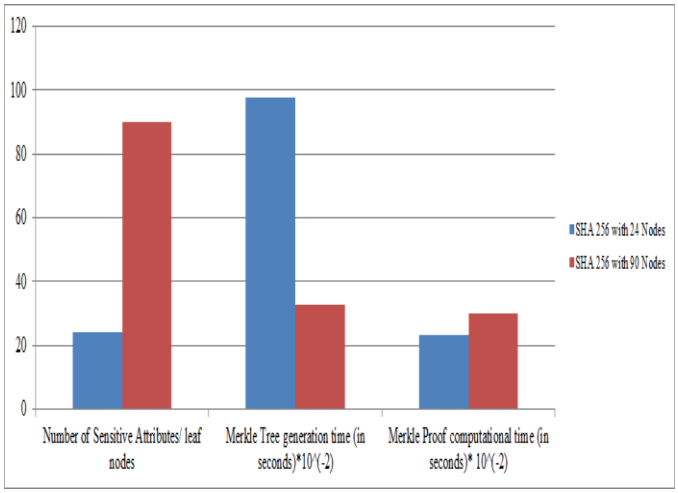

Fig. 5. The sensitive attributes vs tree generation and proof.

TABLE II. Comparison of Proposed Model with other Frameworks

S.No.	Model	CPU Load	Integrity	Efficiency	Proof of work	Privacy
1.	Siti Dhalila et al., [6]	Moderate	Yes	Improved	No	Through fully hiding Full Policy/ file hiding through one-way Hashing
2.	John Bethencourt et al.,[15]	Moderate	No	Nominal	No	The model applicable to large universe of attributes. Policy and Attribute are exposed to unauthorized
3.	Proposed model	Less	Yes	High	Yes	Partial sensitive attribute hiding

2) Replay attack: In order to protect against replay attacks, one approach is to include a timestamp or nonce in the data block that is hashed to create the leaf node of the Merkle tree. When a user requests access to a sensitive attribute, the system will re-compute the hash value of the leaf node using the current timestamp or nonce. This ensures that each request is unique and cannot be replayed at a later time.

For example, consider a user requests access to a sensitive attribute at time t1, the system will compute the hash of the data block with the timestamp t1. If the same user tries to access the same attribute again at time t2, the system will compute the hash of the data block with the timestamp t2, which will be different than the previous hash computed at time t1, hence it will prevent replay attack as the hash values are different.

VI. CONCLUSION

The Merkle tree based access structure in CPABE addresses the attribute disclosure problem in managing access to sensitive patient information by providing fine-grained control and preserving privacy through attribute hiding. The proposed model, a proof-of-work algorithm, enhances security and integrity for both users and providers in healthcare organizations by using a monotonic access structure for policy transmission and being resistant to chosen-ciphertext attacks and replay attacks. It can also handle large volumes of data and provide quick integrity proof. While the model addresses limitations of existing solutions and provides a secure solution for healthcare organizations, further research is needed to explore its potential in addressing other security challenges and improving its efficiency. Overall, this study makes a significant contribution to the field of secure access control for sensitive patient information.

REFERENCES

[1] Maesa, Damiano di Francesco, Paolo Mori and Laura Ricci. “Blockchain Based Access Control.” IFIP International Conference on Distributed Applications and Interoperable Systems (2017).

[2] H. -N. Dai, M. Imran and N. Haider, “Blockchain-Enabled Internet of Medical Things to Combat COVID-19,” in IEEE Internet of Things Magazine, vol. 3, no. 3, pp. 52-57, September 2020.

[3] Hu, Vincent, Tim Grance, David Ferraiolo, and D. Kuhn. “An Access Control Scheme for Big Data Processing”. In Proceedings of the 10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing. ICST, 2014. [4] Huang, Liang, Hyung-Hyo Lee, and Hongju Cheng. “A Medical Data Privacy Protection Scheme Based on Blockchain and Cloud Computing”. Wirel. Commun. Mob. Comput. January 2020.

[4] Yüksel, B., Küpçü, A., & Özkasap, Ö. “Research issues for privacy and security of electronic health services”. Future Gener. Comput. Syst., 68:1–13, 2017.

[5] Barbara Carminati. Merkle Trees. Encyclopedia of Database Systems. 2009.

[6] Siti Dhalila Mohd Satar, Mohamad Afendee Mohamed, Masnida Hussin, Zurina Mohd Hanapi and Siti Dhalila Mohd Satar, “Cloud-based Secure Healthcare Framework by using Enhanced Ciphertext Policy Attribute-Based Encryption Scheme” International Journal of Advanced Computer Science and Applications(IJACSA), 12(6), 2021.

[7] Engelhardt, Mark A.. “Hitching Healthcare to the Chain: An Introduction to Blockchain Technology in the Healthcare Sector.” Technology Innovation Management Review 7 (2017): 22-34.

[8] S. Jiang, J. Cao, H. Wu, Y. Yang, M. Ma and J. He, “BlocHIE: A BLOCkchain-Based Platform for Healthcare Information Exchange,” 2018 IEEE International Conference on Smart Computing (SMARTCOMP), Taormina, Italy, 2018, pp. 49-56.

[9] Dinh C. Nguyen, Pubudu N. Pathirana, Ming Ding, Aruna Seneviratne. “A cooperative architecture of data offloading and sharing for smart healthcare with blockchain”. 2021. IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2021.

[10] Shulan Wang, Haiyan Wang, Jianqiang Li, Huihui Wang, Junaid Chaudhry, Mamoun Alazab, Houbing Song. “A Fast CP-ABE System for Cyber-Physical Security and Privacy in Mobile Healthcare Network. IEEE Transactions on Industry Applications”. 564. 2020.

[11] Chang Ji Wang, Xi Lei Xu, Dong Yuan Shi, Wen Long Lin. “An efficient cloud-based personal health records system using attribute-based encryption and anonymous multi-receiver identity-based encryption”. Proceedings - 2014 9th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing.

[12] Leyou Zhang, Yadi Ye, Yi Mu. “Multiauthority Access Control with Anonymous Authentication for Personal Health Record”. IEEE Internet of Things Journal. 81. 2021.

[13] Amit Sahai, Brent Waters. “Fuzzy identity-based encryption”. Lecture Notes in Computer Science. 3494. 2005.

[14] Vipul Goyal, Omkant Pandey, Amit Sahai, Brent Waters. “Attribute-based encryption for fine-grained access control of encrypted data”. Proceedings of the ACM Conference on Computer and Communications Security. 2006.

[15] John Bethencourt, Amit Sahai, Brent Waters. “Ciphertext-policy attribute-based encryption”. Proceedings - IEEE Symposium on Security and Privacy. 2007.

[16] Hui Cui, Robert H. Deng Junzuo Lai, Xun Yi, Surya Nepal. “An efficient and expressive ciphertext-policy attribute-based encryption scheme with partially hidden access structures, revisited”. Computer Networks. 133. 2018.

[17] Vini Vijayan, James Connolly, Joan Condell, Nigel McKelvey, Philip Gardiner. “Review of wearable devices and data collection considerations for connected health”. Sensors 2116. 2021.

[18] Takashi Nishide, Kazuki Yoneyama, Kazuo Ohta. “Attribute-based encryption with partially hidden encryptor-specified access structures”. Lecture Notes in Computer Science. LNCS. 2008.

[19] “Longitudinal Patient Records Artifacts.” Longitudinal Patient Records Artifacts, www.ibm.com, 12 Apr. 2021.

[20] Nishant Doshi, Devesh Jinwala. “Constant ciphertext length in multi-authority ciphertext policy attribute based encryption”. 2011 2nd International Conference on Computer and Communication Technology, ICCCT-2011.

[21] Nurmamat Helil, Kaysar Rahman. “CP-ABE access control scheme for sensitive data set constraint with hidden access policy and constraint policy”. Security and Communication Networks. 2017.

[22] https://pymerkle.readthedocs.io/en/latest/index.html#

[23] B. Ravinder Reddy, T. Adilakshmi, “Merkle Tree-based Access Structure for Sensitive Attributes in Patient-Centric Data,” International Journal of Engineering Trends and Technology, vol. 70, no. 6, pp. 106-113, 2022.

APPENDIX A:

TABLE III. SYMBOL TABLE

S.No.	Symbol	Description
1.	ABE	Attribute Based Encryption
2.	{a1,a2,a3…an}	Attributes
3.	A,B,C,S	Attribute Set
4.	S_T	Sensitive attribute Access Tree
5.	NS_T	Non-Sensitive attribute Access
6.	Ss_T	Sub Tree
7.	T	Threshold value
8.	Ks	Child node
9.	num_s	Number of children of s
10.	H	Hash function
11.	R_H	Root Hash
12.	λ,β	Security Variable
13.	MS_K	Master Key
15.	S_K	Secret Key
16.	K_G	Key generation
17.	K_{GC}	Key Generation Centre
18.	PCD	Patient-Centric Data
19.	M	Input Message
20.	HL7	Health Level 7
21.	C_T	Cipher Text
22.	E	Encryption
23.	D	Decryption