A PRIVACY PRESERVING MEDICAL DATA SHARING SCHEME BASED ON BLOCKCHAIN

ADDEPALLI ARUNA[1]

CHANDU DELHI POLICE[2]

[1]PG scholar, [2]Associate professor, Department of Computer Science Engineering Priyadarshini Institute of Technology & Science, Tenali, Guntur

Abstract: The limits of conventional medical systems are becoming apparent as the Internet of Things (IoT) becomes more and more integrated into people's lives. First, privacy exposure might readily result from the standard method of managing sensitive information. Secondly, the health care system is somewhat remote. Within the boundaries of a medical system, users' actions are restricted and data sharing across systems is challenging. By integrating the authorization mechanism and attribute-based encryption (ABE) based on blockchain, we offer a novel privacy-preserving medical data-sharing scheme that addresses these two issues. This scheme breaks down system borders and realises data sharing across different medical institutions. Scalable access control is achieved with the help of ABE. Additionally, by implementing many-to-many matching, which enables patients' health data to be represented by numerous keywords and physicians' expertise to be represented by various interests, doctors may share their knowledge to diagnose users. We provide the security and correctness analysis of our plan and put our Ethereum prototype tool into use. The outcomes of the trial demonstrate that our plan resolves the conflict between the need for data exchange and the privacy protection of medical records.

Index Terms: Attribute-based encryption, blockchain, privacy preserving, intelligent medical system.

I. Introduction

Advances in mobile communication and the internet have promoted the popularization of the Internet of Things (IoT). IoT connects many hardware devices and realizes data sharing on these devices to help improve monitoring and management. IoT has been widely used in smart transportation, smart retail, smart agriculture and other fields. For example, the agriculture-related data like soil properties and water level can be monitored in real time through IoT devices. Through a privacy-preserving data aggregation scheme based on ElGamal Cryptosystem in [51], we can make a balance between data sharing and privacy preservation. IoT has also made great achievements in the intelligent medical scenario. In an intelligent medical system, patients' physiological data are captured by sensors, processed centrally in the local gateway, and then sent to the medical service provider. Doctors can obtain patients' health data anywhere through the intelligent medical technology to facilitate remote medical treatment. Such an intelligent system realizes the interconnection between patients and doctors.

Although the intelligent medical system has changed the traditional medical treatment process and dissolved geographical restrictions associated with the traditional process, the existing intelligent medical system can only work independently and lacks cooperation. Patients can only initiate consultation within the system boundary. For example, patients who are in some specialized hospitals cannot be treated for diseases requiring other specialties, and some hospitals are not qualified enough to diagnose patients with more serious diseases. Township and community hospitals cannot achieve seamless information connection with large hospitals. This inability greatly limits the quality of medical treatment. To resolve it, we developed a new privacy-preserving medical data sharing scheme that can interconnect decentralized medical service providers to form a joint platform on the premise of mutual authorization between hospitals. It is easy to obtain expert opinions in real time and achieve data sharing when transferring from one hospital to another.

Blockchain can be regarded as a distributed recording ledger with the characteristics of anonymity, tamperability, auditability and autonomy. The smart contract (SC) running on it can avoid the interference of malicious users in the normal operation process, which is an effective solution for privacy-preserving medical data sharing. There are still many practical challenges to be solved when applying the blockchain to medical data sharing. The following are the major issues that interest us. 1) Patient medical data are highly sensitive and should be protected when uploading to the blockchain. Doctors and hospitals may compromise the patient's privacy without their permission for commercial interests. 2) Information sharing between medical institutions contradicts the high sensitivity of patient medical data.

In this paper, we construct a new privacy-preserving data-sharing scheme based on the blockchain for medical scenarios, which breaks system boundaries and realizes data sharing among several medical institutions. Patients can upload their own electronic medical records in privacy, the authorized relevant hospitals can conduct private keyword matching, the matched doctors can diagnose patients based on the provided electronic medical records, and the diagnosis results can be safely transmitted to the corresponding patients. Because the professional field of doctors has personalized attributes, this paper adopts attribute-based encryption to grant the data owner (i.e., the patient) the right to control the access of data. The matched doctors must decrypt the health data according to their own attributes before diagnosis. In addition, we introduce an authorized and revocable mechanism to ensure that doctors in authorized hospitals can obtain access to the data while the revoked doctor cannot obtain. Moreover, we use a zero-knowledge proof protocol to ensure the credibility of the hospital matching algorithm under privacy conditions.

II. Literature Review

As a database, the blockchain has been widely used in the field of decentralized networks [45], [56]. Blockchain technology has many characteristics [32]. The blockchain with decentralization and anonymity has become the core technology behind Bitcoin. In addition to its application in the financial field, the blockchain has also been applied to other fields, including smart transportation [29], [58], smart grid [31], and data auditing [25]– [27].

In the field of smart grid, malicious users may infer a user's private information from electricity consumption data. For this, Guan et al. introduced a data aggregation scheme based on blockchain to preserve users' privacy [31]. Users' identities are hidden in pseudonyms. A user can be associated with multiple pseudonyms to submit electricity consumption data. And in intelligent transportation system, Ning et al. constructed a crowdsensing framework based on the decentralized blockchain to realize key management in a distributed way [29], which makes a trade-off between minimizing the latency and maximizing the safety of the blockchain. Many scholars are committed to using blockchain to improve the security of Unmanned Aerial Vehicles (UAV). Ch et al. [55] presented a Blockchain Technology (BCT) based solution to better manage sensitive data and prevent data from being attacked.

However, the application of these schemes often requires specific scenarios, so it is inappropriate to apply them directly to intelligent medical scenarios. Both the patient's health data and the health report generated by the doctor should be protected. Therefore, for medical scenarios, we propose a scheme based on the blockchain that can realize data sharing on the premise of protecting user privacy.

In this era of information explosion, data means resources, but the more such resources, the greater the possibility of privacy leakage. People try to illegally occupy and exploit resources, and the privacy security has not been ignored. Many scholars are committed to enhancing the privacy protection and reducing illegal appropriation of resources [1], [16]– [18], [57]. The intelligent medical system is no exception. The transformation from the traditional health care system to the electronic health care system has made clinical data easier and faster to access. However, inevitably, patients' privacy concerns cannot be ignored. The patient's body data are personal and sensitive. Direct exposure to the shared cloud environment will eliminate data privacy [3], which will not only affect relevant laws and regulations but also have a serious economic and social impact on the patient. Therefore, it is an urgent problem to introduce a strong privacy protection mechanism into the whole intelligent medical system. A mechanism to reduce the linkability between patients and medical records is proposed by Li et al. [7]. Hupperich et al. [10] mentioned out that the existing privacy protection is either too strict and requires patients to be available to authorize access to medical records or is insufficient and does not truly realize privacy protection. Therefore, it is necessary to provide more flexibility for whole system to ensure that doctors can access medical records without the presence of patients. Abbas et al. [4] mentioned out that the privacy preserving methods commonly used in intelligent medicine are divided into two categories: 1) cryptographic approaches, which use specific encryption primitives to reduce privacy risks, and 2) noncryptographic approaches, which mainly adopt a policy-based infrastructure, to standardize the access control of data. A new access control mechanism is provided [11], which is a noncryptographic approach to support the fine-grained sharing of electronic health records from different medical service providers in the cloud by implementing the access control policies specified by patients. Cui et al. [4] combined attribute-based encryption with keyword search in a cloud storage system that keeps personal health records. However, it requires the cloud server to be honest, and this scheme only consider the scene where patients are treated for diseases in a single medical institution rather than multiple institutions. This limitation is not compatible with the requirement to break the geographical restrictions of medical institutions.

III. Proposed system

In this section, we introduce the basic mechanism of attribute based encryption (ABE) in detail. ABE is a relatively new encryption mechanism which does not require information interchange between data owners and users and is suitable for the one-to-many distribution. The user's private key and ciphertext are constructed based on the attribute set and access policy. As long as the attribute set matches the policy, plaintext can be obtained. Many studies have sought to optimize the algorithm and improve the efficiency. Jin et al. [19]

proposed a secure and lightweight data access control scheme. Most of the computation operations are performed by the cloud and the computing overhead of users is greatly reduced. Lewko et al. [20] designed two functional encryption schemes including an attribute-based encryption and a predicate encryption for inner product predicates, which are fully secure systems. Xu et al. [2] proposed a privacy-enhanced access control mechanism. Different schemes are implemented according to whether the attribute belongs to the sensitive attributes, and finally the comprehensive decision is made. Sensitive data should not be uploaded directly before outsourcing but should be encrypted; otherwise, user privacy will be compromised. However, the encrypted data affect the efficiency of data querying [21]. Searchable encryption allows users to search directly on ciphertext using keywords without decrypting data [22], which solves the contradiction between data confidentiality and searchability. In searchable encryption, users are often allowed to generate a trapdoor according to their own interests and then match the trapdoor with keywords. Attribute-based searchable encryption successfully combines the advantages of the two. Due to the high computational cost of anonymity, Han et al. introduced a weak anonymity feature to hide users' identity. Based on this weak anonymity, they proposed a general transformation from attributebased encryption to attribute-based encryption with keyword search and constructed a specific attribute private key-policy ABE scheme, which allows multiple users to flexibly search remote encrypted data [23]. However, most existing attribute-based searchable encryption approaches support only one-to-one or one-to-many retrieval, and few approaches support many-to-many retrieval. In our design, we realize many-to-many retrieval based on attribute-based encryption.

As shown in Fig. 1, our model involves five entities: patients, doctors, hospitals, the key generation center (KGC), and blockchain platforms. Symbols used in this scheme are shown in Table II. 1) Patients: As the requesters of medical consultation system services, patients can upload their own physiological data or medical records to the hospital. Patients can obtain the information from the hospital to know their physical status.

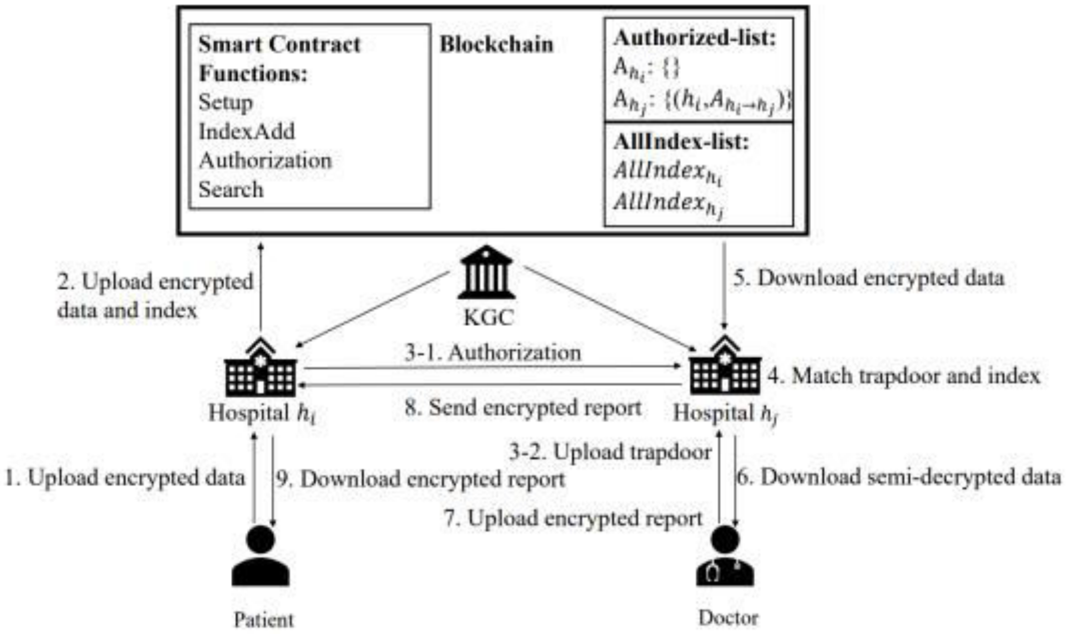


Figure 1: The data flow of the system.

2) Doctors: Doctors can match the data uploaded by patients according to their own interests, complete the diagnosis after successful matching, and transmit the diagnosis results to the corresponding patients. 3) Hospitals: All patients and doctors belong to a hospital. Each hospital has corresponding patients and doctors. Doctors provide services to patients through hospitals. As service providers, hospitals are responsible for matching patients and doctors. 4) KGC: The KGC is honest and generate the predecryption key according to the public key and the attributes of the doctor. The leaving doctor will be added to the revocation list by the KGC. 5) Blockchain platform: Blockchain and smart contracts connect all hospitals to form a joint medical consultation system. All hospitals share data through the blockchain.

B. System Procedure

The data flow in this scheme is shown in Fig. 1. The procedure of the whole scheme is as follows.

1) System Initialization: The whole system is initialized. Setup(1λ) → (pp, msk, rl, st) is run by the KGC. Given a security parameter λ, the public parameter pp, the master private key msk, a revocation list rl, and a system state st are output by the algorithm.

2) User Registration and Revocation: Anyone who wants to join the system needs to run UserKG(pp, id) → (skid, pkid). The algorithm generates its identity key pair according to the

id of the user. Revoke(id, tm, rl, st) → rl will be run to revoke the user by updating the revocation list. The algorithm takes the id of the

revoked user, time mark tm, and the system state st as inputs and a new revocation list is output. 3) Encryption: If a user wants to upload his or her health data, the data must be encrypted by running a symmetric algorithm using SymKey and the data owner can obtain CTHD. Then, Encrypt(pp,(M, ρ), tm, SymKey) → CTSym is run by the data owner to hide the Symkey. Then, the CTHD and CTSym will be sent to hospital hi , where the patient visits.

4) Index Generation: IndexBuild(pp, {H(kw)}) → Index is run by the hospital where the patient wants to be treated. kw is a keywords set selected by the patient from the data, and H is a hash function. Then, the index is created and will be uploaded to the blockchain.

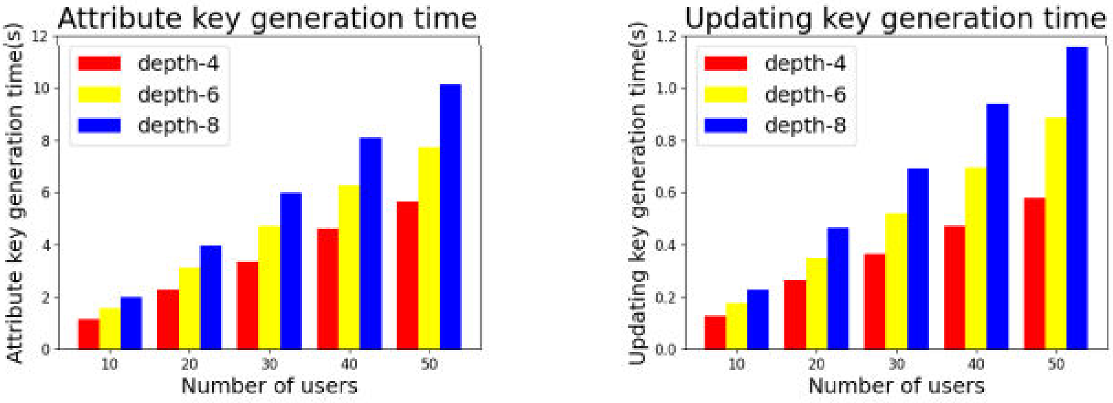
5) Authorization: If a hospital does not have enough doctors or the patients' disease exceeds the hospital's diagnostic ability, the hospital will authorize access to the patient's data to other hospitals that have the ability to diagnose the patient. The hospital will generate Ahi→hj , which means hi authorizes to hj and publishes it to the blockchain. 6) Search: The authorized hospital can obtain the correct index using the provided Ahi→hj . 7) Key Generation: There are three steps to complete. First, AttKeyGen(pp, msk, id, pkid, Aid, st) → (attkid, st) is run by KGC according to the patient's attributes Aid. Then, U pdKeyGen(pp, msk, tm, rl, st) → (updktm, st) is run by KGC according to the revocation list to ensure that the revoked doctor cannot obtain the key updktm. P reKeyGen(pp, id, attkid, updktm) → prekid,tm is finally executed taking the attkid and updktm generated in the first two steps as input. The prekid,tm will be used to predecrypt. 8) Trapdoor Generation: If a doctor wants to obtain the patients' data, he or she needs to run T rdGen(pp, {H(inst)}, m1) → T rapdoor. This step takes the hash of the doctor's interest inst and the number of keywords m1 as inputs. Note that a doctor can correspond to multiple interests. Then T rapdoor will be sent to the hospital where the doctor works for matching. 9) Match: Match(T rapdoor, Index) → Addr/⊥ is run by the hospital to help doctors find the index corresponding to their interest. If the condition that the doctors' interest set is a subset of the keyword set is satisfied, the match is successful. If the match succeeds, it returns the address of the data. Otherwise, it returns ⊥. If multiple doctors matched the same set of keywords, the doctor with more interests is selected to complete the diagnosis. 10) Decryption: Two steps are required to obtain the plaintext of SymKey. First, P reDec(pp, id, Aid, prekid,tm, CTSym, tm) → SemiCT /⊥ is run

by the hospital to transform CTSym to SemiCT. Then the doctor runs Dec(pp, skid, SemiCT) → SymKey/⊥. If the attributes of the doctor meet the requirement, the symmetric key can be decrypted. Using SymKey and CTHD, the doctor can obtain the plaintext of the health data.

11) Results Return: The doctor can then diagnose the patient. The health report is encrypted to CTHR using the RSA encryption algorithm [46]. Finally, CTHR is sent to the patient by hospitals. The patient can decrypt it using the private key to obtain the health report.

IV. Results and Discussion

To access the performance of the scheme we designed, we implemented the prototype design in Python and built a smart contract on Ethereum with Solidity using approximately 2000 lines of code. We simulate in an Ubuntu 20.04 desktop system with Intel Core i7 and generation time of the attribute key attk, updating key updk, and predecryption key prek. From Fig. 2(a), we know that the attribute key generation time varies as the depth of the binary tree deepens. The greater the depth of the tree is, the more users the tree can carry and the more time it takes to generate the attribute key.



(a) (b)

Figure 2: Execution time under varying depths of the binary tree and varying revocation list lengths. (a) Attribute key generation time under varying depths. (b) Updating key generation time under varying depths.

The change of updating key generation time is similar to the attribute key as shown in Fig. 2(b). Different from Fig. 2(a) and Fig. 2(b), the depth of the binary tree has no significant effect on predecryption key generation time. We agree that the best explanation for this result is that in the predecryption key generation stage, we select the elements in the intersection of sets P ath(θ) and KUNodes(BT, rl, tm) for operation. Although the number of elements in set I and set J is affected by the depth of the tree, the time taken to obtain the intersection of the two sets is negligible.

V. Conclusion

We suggested a novel blockchain-based medical data exchange programme that protects patient privacy. Our plan enables many-to-many matching, lightweight decryption, and fine-grained data access. In particular, we encrypt patient medical data using attribute-based encryption technology. Information exchange between different medical institutions may be achieved using the permission process between medical institutions and broadcasting the authorization results to the blockchain. Unauthorised medical institutions are unable to access the patient's health information, whereas doctors at accredited hospitals are able to finish diagnosing patients in hospitals where the patient registers. Physicians may use searches carried out by medical institutions to get data tailored to their interests and areas of competence in order to receive individualised care. In order to prevent the hospital from matching patient data at will owing to its commercial interests, only physicians whose characteristics comply with the criteria of the access policy may successfully decrypt the data indicating that the match algorithm is truly finished with the right index and trapdoor. To stop the malevolent patient from framing physicians and generating medical disputes, the patient must provide a zero knowledge proof demonstrating that the decryption process was successfully completed using the private key that corresponds to the public key after they acquire the ciphertext of the health report. We use the pre-decryption approach for decrypting, and the medical institutions shoulder some of the processing load, in light of the possibility that physicians' local computer capacity may be constrained. To extract plaintext, just one exponentiation operation and one product operation are needed. We offered security and correctness analyses, and we used theoretical and experimental analysis to assess our scheme's effectiveness. Based on the principle of safeguarding medical privacy, our study demonstrates that the suggested scheme is effective, accurate, well-suited to medical circumstances, and capable of realising medical data exchange and enhancing the use of social medical resources.

In spite of our research's advancements, a number of issues still need to be taken into account in the future. The collaborative medical consultant system is still a mostly uncharted territory. Currently, there is a lack of maturity in the creation of smart contracts, which limits the realisation of certain functionalities. Furthermore, the chain often has a limited quantity of data storage. To decrease the amount of data on the chain, we may further optimise our plan, implement more intricate functions, and limit storage on the chain. Secondly, we may extend our implementation via a more comprehensive incentive system. The patient will have to pay a set amount for the doctor's competence after we implement the payment feature in our plan. Third, in order to standardise the whole system, certain rules must be established. Not to mention, our plan relies on a reliable authority to finish key distribution and generation. Creating a way that does not need a reliable third party is a superior alternative. Smart contracts, for instance, may be designed to take the role of reliable, centralised KGC [50]. To increase the scheme's viability, we may also recreate the secret from the consensus node using the secret sharing approach.

VI. References

[1] Y. Xu, X. Yan, Y. Wu, Y. Hu, W. Liang, and J. Zhang, “Hierarchical bidirectional RNN for safety-enhanced B5G heterogeneous networks,” IEEE Trans. Netw. Sci. Eng, vol. 8, no. 4, pp. 2946-2957, 2021.

[2] Y. Xu, Q. Zeng, G. Wang, C. Zhang, J. Ren, and Y. Zhang, “An efficient privacy-enhanced attribute-based access control mechanism,” IEEE Trans. Netw. Sci. Eng., vol. 32, no. 5, pp. e5556, 2022.

[3] X. Yuan, X. Wang, C. Wang, J. Weng, and K. Ren, “Enabling secure and fast indexing for privacy-assured healthcare monitoring via compressive sensing,” IEEE Trans. Multimedia, vol. 18, no. 10, pp. 2002-2014, 2016.

[4] A. Abbas and S. U. Khan, “A review on the state-of-the-art privacypreserving approaches in the e-health clouds,” IEEE JBHI, vol. 18, no. 4, pp. 1431-1441, 2014.

[5] C. Esposito, A. D. Santis, G. Tortora, H. Chang, and K. R. Choo, “Blockchain: A panacea for healthcare cloud-based data security and privacy?” IEEE Cloud Comput., vol. 5, no. 1, pp. 31-37, Jan./Feb. 2018.

[6] S. Haas, S. Wohlgemuth, I. Echizen, N. Sonehara, and G. Muller, ¨ “Aspects of privacy for electronic health records,” Int. J. Med. Informat, vol. 80, no. 2, pp. e26-e31, 2011.

[7] Z. R. Li, E. C. Chang, K. H. Huang, and F. Lai, “A secure electronic medical record sharing mechanism in the cloud computing platform,” in Proc. 15th IEEE Int. Sympo. Consum. Electron., pp. 98-103, Jun. 2011.

[8] Y. Y. Chen, J. C. Lu, and J. K. Jan, “A secure EHR system based on hybrid clouds,” J. Med. Syst., vol. 36, no. 5, pp. 3375-3384, 2012.

[9] S. Yu, C. Wang, K. Ren, and W. Lou, “Achieving secure, scalable, and fine-grained data access control in cloud computing,” in Proc. IEEE Infocom, pp. 1-9, Mar. 2010.

[10] T. Hupperich, H. Lohr, A. R. Sadeghi, and M. Winandy, “Flexible ¨ patient-controlled security for electronic health records,” in Proc. 2nd ACM SIGHT Sympo. Int. Health Informatics, pp. 727-732, Jan. 2012.

[11] R. Wu, G. L. Ahn, and H. Hu, “Secure sharing of electronic health records in clouds,” in Proc. 8th IEEE Int. Conf. Collaborative Comput., Netw., Appl. Work-sharing, pp. 711-718, Jan. 2012.

[12] X. Yue, H. Wang, D. Jin, M. Li, and W. Jiang, “Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control,” J. Med. Syst., vol. 4, no. 10, pp. 1-8, 2016.

[13] B. Shen, J. Guo, and Y. Yang, “MedChain: Efficient Healthcare Data Sharing via Blockchain,” Appl. Sci., vol. 9, no. 6, pp. 1207, 2019.

[14] H. L. Pham, T. H. Tran, and Y. Nakashima, “A secure remote healthcare system for hospital using blockchain smart contract,” IEEE GC. Wkshps., vol. 9, no. 6, pp. 1-6, Feb. 2018.

[15] R. Zou, X. Lv, and J. Zhao, “SPChain: Blockchain-based medical data sharing and privacy-preserving eHealth system,” IPM, vol. 58, no. 4, pp. 102604, 2021.

[16] Z. Lian, W. Wang, H. Huang, and C. Su, “Layer-Based CommunicationEfficient Federated Learning with Privacy Preservation,” IEICE Trans Inf Syst, vol. 105, no. 2, pp. 256-263, 2022.

[17] J. Song, W. Wang, T. R. Gadekallu, J. Cao, and Y. Liu, “Eppda: An efficient privacy-preserving data aggregation federated learning scheme,” IEEE Trans. Netw. Sci. Eng., 2022.

[18] Z. Lian, Q. Yang, W. Wang, Q. Zeng, M. Alazab, H. Zhao, and C. Su, “DEEP-FEL: Decentralized, Efficient and Privacy-Enhanced Federated Edge Learning for Healthcare Cyber Physical Systems,” IEEE Trans. Netw. Sci. Eng., 2022.

[19] Y. Jin, C. Tian, H. He, and F. Wang, “A secure and lightweight data access control scheme for mobile cloud computing,” in Proc. 5th Int. Conf. Big Data Cloud Comput., vol. 15, pp. 172-179, Aug. 2015.

[20] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, “Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption,” EUROCRYPT, pp. 62-91, 2010.

[21] C. Wang, W. Li, Y. Li, X. Xu, “A ciphertext-policy attribute-based encryption scheme supporting keyword search function,” CSS, pp. 377- 386, 2013.