Improving Privacy and Efficiency for Blockchain-based Educational Credential Sharing

***Abstract*—Electronic educational credential sharing is a typical process involving two or more independent institutes. To address the trust problem among different institutes when sharing educational credentials, blockchain technology has recently been adopted to establish a decentralized data-sharing system. However, storing educational credentials on the blockchain suffers from privacy issues since the blockchain can be accessed by all nodes participating in the decentralized system. This paper addresses the privacy issue of blockchain-based educational credential sharing via searchable encryption and proxy re-encryption. By employing proxy re-encryption and searchable encryption, the confidentiality of students' credentials in the blockchain can be guaranteed, and meanwhile, authorized searching is allowed. In addition, a reputation-based consensus mechanism is designed to improve the efficiency of the system and motivate institutes to join the credential sharing system.**

***Index Terms*—Educational Credential Sharing, Blockchain, Data Sharing, Proxy Re-Encryption, Searchable Encryption, Reputation Mechanism**

I. Introduction

Educational credentials are significant indicators of students' abilities and qualifications for higher education or employment, so universities and companies demand high standards of authenticity for students' credentials. To deter credential tampering, universities often add a unique number and detailed personal information. However, this poses two main challenges: authenticity and privacy.

Cuurrently, many universities accelerate the issuance and validation of educational credentials through digitization. They use centralized servers to store and manage students' credentials, and online sharing of credentials via email to ensure authenticity. Such guarantees are difficult, even with such complex certificate sharing processes, there is still the potential for authenticity violations. Malicious institutions may build fake servers and issue fake certificates to get profit [1, 2].

The authenticity problem can be addressed by blockchain through which a decentralized and tamper-proof ledger is maintained by multiple stakeholders. Blockchain technology can be found in an abundance of literature and has been implemented for educational applications, e.g. credit transfer [3]. Blockchain could also provide a safe and reliable record keeping of students in the architecture of Arcinas [4]. Some studies have reported methods for blockchain on education credential sharing, but it is in its infancy. Gräther et al. [5] introduced a blockchain-based open education platform to issue, validate, and share certificates, and the certificate information is stored in the interplanetary file system (IPFS). Ayub Khan et al. [6] proposed a detailed design of blockchain-based hyperledger fabric applications for degree attestation verification. However, the architecture of the above papers all stores the plaintexts of credentials in the blockchain, which may lead to privacy leakage.

To protect the transaction data in the blockchain, many researchers emphasize the integration of blockchain with privacy-preserving algorithms. Li and Han [7] implemented a consortium blockchain-based system for sharing educational records and designed three smart contracts to execute the access control of record sharing, but it ignored the incentive for education institutions. Liu et al. [8] proposed a two-stage blockchain to provide a trustworthy authentication and storage of educational information, where zero-knowledge proofs are used to preserve sensitive data, the Vickrey-Clerke-Groves game is used to find the Nash Equilibrium. Li and Ma [9] implemented a system for sharing educational data based on a consortium blockchain and searchable encryption algorithms. The students' sensitive data are stored on a cloud server, and the smart contract validates the requester's access rights. Mishra et al. [10] developed an Ethereum-based architecture using off-chain distributed file storage to increase scalability and lower system operating costs and put their digital fingerprints on the blockchain for sharing students' credentials in the education ecosystem. This could be a cumbersome task as their scheme's search is based on a unique ID, which requires a government department to create a unique lifetime account for each user. Hou et al. [11] developed a privacy-preserving blockchain-based educational credentials sharing scheme by proxy re-encryption to improve confidentiality.

To sum up, it is challenging to protect the students' sensitive information and meanwhile provide educational credential sharing convenience and searchability since blockchain is public and transparent. Besides, the common consensus mechanisms, such as PoW and PoS, have low efficiency in block generation. This leads to a high delay in uploading data to the blockchain for sharing. To address the privacy and efficiency challenges faced by blockchain-based educational credential sharing, the paper employs proxy re-encryption and searchable encryption techniques to achieve privacy protection and proposes a novel consensus mechanism to improve efficiency. In particular, keyword ciphertext stored in the blockchain ensure that requesters can match the intended students, and proxy re-encryption ensures that only authorized requesters can access the student's credentials. The new consensus mechanism reduces consensus time and stimulates institutes' contributions by linking the power of generating new blocks with reputation calculated from contributions.

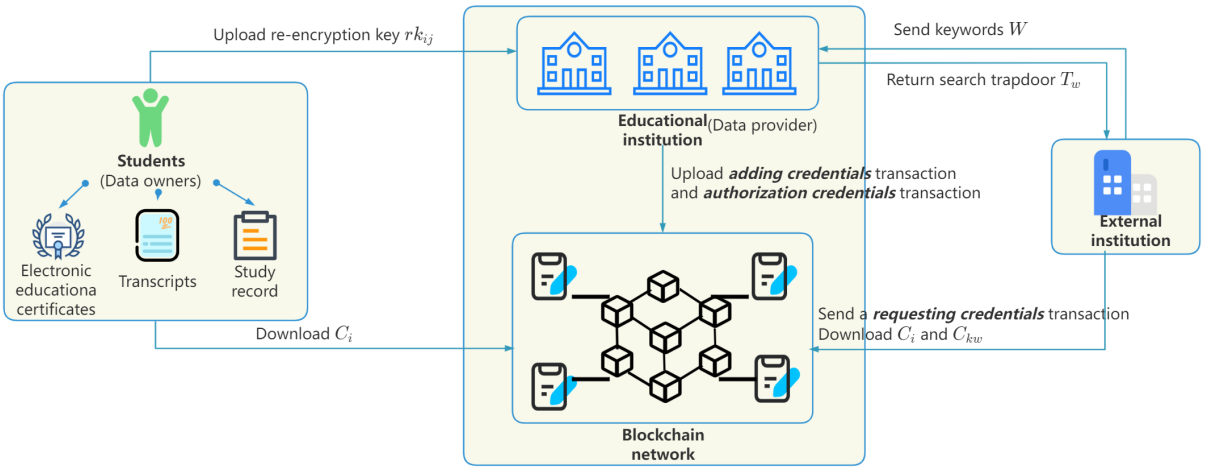

Fig. 1. System overview.

II. System Model

In this section, the system architecture and blockchain architecture of the educational credential sharing scheme are illustrated, respectively.

A. System Architecture

The proposed system consists of four entities: educational institution (data provider), student (data owner), and external institution (data requester). The overview of system architecture is shown in Fig 1. More details of the system are as follows.

1) Student: The student is the data owner who comes to the educational institution. After the student finishes the courses or degrees from the educational institution and obtains electronic educational certificates, transcripts, study records, etc., the educational institution uploads the ciphertext of educational credentials to the blockchain, and then the student can download the ciphertext from the blockchain and use his or her private key to decrypt the ciphertext. Once receiving the request of an external institution, the student performs an operation and sends a re-encryption key to the educational institution to permit the external institution to access the educational credentials.

2) Educational Institution.: The administrator of the educational institution is the data provider. After a student finishes the courses or degrees, the educational institution performs the encryption operation to encrypt the student's credentials by the student's public key, and encrypt the keywords of the credentials by the educational institution public key, and then uploads the ciphertext of credentials and keyword to the blockchain. Upon receiving the student's re-encryption key, the educational institution runs a re-encryption algorithm to generate a re-encryption ciphertext and uploads it to the blockchain.

3) External Institution.: The external institution, such as an enterprise or a university, is the data requester that needs the student's access authorization. It can get a search trapdoor from the educational institution and obtain the search results by running a search algorithm. The requester uploads a request transaction to the blockchain if it needs to view a student's credential. Once the requester is authorized by the students, it can download a re-encryption ciphertext from the blockchain and decrypt the ciphertext by its private key.

B. Blockchain Architecture

This consortium blockchain consists of two types of nodes: committee nodes (educational institutions) and client nodes (students and requesters). Each commissioner node maintains the blockchain state and has the right to upload the credential transaction, and each client node can download the ciphertext from the blockchain. The design of the consensus mechanism and smart contract is described as follows.

1) Consensus Mechanism.: Delegated Proof-of-Stake (DPoS) [12] is an efficient, decentralized, and flexible consensus mechanism. It elects 21-101 manager nodes through a fair and democratic voting process, and only manager nodes are responsible for participating in the consensus mechanism, which can reduce the computational cost and speed up the speed of block generation. In our system, the education institutions with higher reputations are selected as managers nodes of the consortium blockchain, and the reputation mechanism is described in Section II-B3. One of the manager nodes, as the proposer, is required to pack transactions and generate blocks, and other manager nodes should validate the block. Only if the block verification is successful, the current block is linked to the blockchain. At the beginning of the next round, a set of manager nodes is re-selected based on the reputation of the previous round of nodes.

2) Smart Contracts.: To implement the educational credential sharing scheme, the smart contract realizes three functions to be invoked through three transactions, which are the *adding credentials* transaction, *requesting credentials* transaction, and *authorization credentials* transaction.

- *Adding credentials* transaction is sent by the educational institution. It includes the ciphertext of the student's education credentials and keywords.
- *Requesting credentials* transaction is sent by the requester. It includes the students' blockchain address and the public key of the requester.
- *Authorization credentials* transaction is sent by the educational institution. It includes a re-encryption ciphertext of the requester.

3) Reputation Mechanism: The reputation of the educational institution is calculated based on the number of transactions uploaded by the educational institution. The reputation rule is shown in the following formula.

$$\text{Rep}(T) = \text{Rep}(T-1) + 2 \cdot N_{add} + 1 \cdot N_{auth} \tag{1}$$

where $\text{Rep}(T)$ represents that reputation of educational institution in T time; N_{add} and N_{auth} represents the number of *adding credentials* transaction and *authorization credentials* transaction, respectively. In this proxy re-encryption algorithm, the encryption algorithm consumes more computational cost than the re-encryption algorithm, so the *adding credentials* transaction generates more reputation or increases reputation value than *authorization credentials* transaction. The manager nodes in the consortium blockchain are selected based on reputation value.

III. System Design

This section describes the implementation of the proposed system in detail.

A. Education Credential Sharing Process

This education credential sharing scheme realizes three functions: credentials storage, credential searching, and credential sharing.

1) System Initialization: Given a security parameter λ, the committee of blockchain generates a prime number q and a bilinear pairing $e : G_1 \times G_1 \to G_2$. Six one-way collision-resistant hash functions are selected $H_1 : \{0,1\}^l \to Z_q^*$, $H_2 : G_1 \to \{0,1\}^l$, $H_3 : \{0,1\}^* \to Z_q^*$, $H_4 : G_1 \to Z_q^*$, $H_5 : \{0,1\}^* \to G_1$, and $H_6 : G_2 \to \{0,1\}^{log\ q}$. Finally, the system parameters are published as $params = (q, g, G_1, G_2, H_1, H_2, H_3, H_4, H_5, H_6)$.

In order to manage the blockchain, each educational institution randomly selects a private key $sk_{edu}, sk_{edu} \in Z_q^*$ and calculates a public key $pk_{edu} = g^{sk_{edu}}$. In addition, students and requesters run the $KeyGen$ algorithm of encryption based on the parameters $params$ to get their public-private key pairs. Students randomly select the private key $sk_i = (sk_{1i}, sk_{2i}), sk_{1i} \in Z_q^*, sk_{2i} \in Z_q^*$ and calculate the public key $pk_i = (pk_{1i}, pk_{2i})$ by Eq. 2. The requesters randomly select the private key $sk_j = (sk_{1j}, sk_{2j}), sk_{1j} \in Z_q^*, sk_{2j} \in Z_q^*$ and calculate the public key $pk_j = (pk_{1j}, pk_{2j})$ by Eq. 3.

$$pk_{1i} = g^{sk_{1i}},\ pk_{2i} = g^{sk_{2i}} \tag{2}$$

$$pk_{1j} = g^{sk_{1j}},\ pk_{2j} = g^{sk_{2j}} \tag{3}$$

2) Data Uploading: After students complete their courses or degrees, the educational institution generates an electronic educational credential M, $M \in \{0,1\}^l$. The educational institution generates a random number u, and encrypts the credentials M to the ciphertext $C_i = (D, E, F, s)$ using the student's public key (pk_{1i}, pk_{2i}) by Eq 4-7.

$$D = (pk_{i1}^{H_4(pk_{i2})} + pk_{i2})^u \tag{4}$$

$$E = (pk_{i1}^{H_4(pk_{i2})} + pk_{i2})^{H_1(M)} \tag{5}$$

$$F = H_2(g^{H_1(M)}) \oplus M \tag{6}$$

$$s = u + H_1(M) \cdot H_3(D, E, F) \bmod q \tag{7}$$

In order to realize the searching function, the educational institution selects a set of keywords $KW = (kw_1, kw_2, ..., kw_n)$ based on the student's educational credentials. It computes the ciphertext of searchable encryption $C_m, m \in \{1, ..., n\}$ and outputs $C_{kw} = (C_1, C_2, ..., C_n)$.

$$t_m = e(H_5(kw_m), pk_{edu}^r) \tag{8}$$

$$C_m = (g^r, H_6(t_m)) \tag{9}$$

The educational institution sends the ciphertext, including C_{kw} and C_i, to the blockchain as a *adding credentials* transaction.

The student can download the ciphertext of educational credentials C_i from blockchain and decrypt it by own private key (sk_{1i}, sk_{2i}) to get the credential M.

$$\text{assert}((pk_{i1}^{H_4(pk_{i2})} + pk_{i2})^s == D \cdot E^{H_3(D,E,F)}) \tag{10}$$

$$M = F \oplus H_2(E^{\frac{1}{sk_{i1} \cdot H_4(pk_{i2}) + sk_{i2}}}) \tag{11}$$

$$\tag{12}$$

3) Data Searching: The requester sets the customized search keywords $W = (w_1, w_2, ..., w_n)$ and send W to the education institution. The education institution calculates a search trapdoor $T_m, m \in \{1, ..., n\}$ and returns these trapdoors $T_w = (T_1, T_2, ..., T_n)$ to the requester.

$$T_m = H_5(w_m)^{sk_{edu}} \tag{13}$$

The requester downloads the ciphertext of educational credentials keywords C_w from the blockchain, and matches the keywords by the search operation Eq. 14. The algorithm returns $true$ if the student's educational credentials meet the requester's keyword criterion; otherwise, the search algorithm returns $false$.

$$\text{assert}(H_6(e(T_n, g^r)) == C_n) \to \text{true/false} \tag{14}$$

4) Data Sharing: The requester sends a *requesting credentials* transaction to the blockchain if it wants to view the student's credentials. After being notified of the access request of the requester, the student generates a random number h and

calculates a re-encryption key rk_{ij} by Eq. 15-17, and then sends the key to the educational institution.

$$V = pk_{j2}^{v}, v = H_1(h) \tag{15}$$

$$W = H_2(g^v) \oplus (h) \tag{16}$$

$$rk_{ij} = \frac{h}{sk_{i1} \cdot H_4(pk_{i2}) + sk_{i2}} \tag{17}$$

The educational institution calculates a re-encryption ciphertext $C_j = (E', F, V, W)$ in Eq. 18, and sends C_j as an *authorization credentials* to the blockchain.

$$E' = E^{rk_{ij}} \tag{18}$$

The requester downloads the re-encryption ciphertext $C_j = (E', F, V, W)$ from blockchain and decrypts it using own private key (sk_{j1}, sk_{j2}) by Eq. 19-20. The requester can view the student's educational credentials M.

$$h = W \oplus H_2(V^{\frac{1}{sk_{j2}}}) \tag{19}$$

$$M = F \oplus H_2(E'^{\frac{1}{h}}) \tag{20}$$

B. Consensus Mechanism Process

The improved DPoS consensus mechanism uses the voting method to perform block verification. The block is generated if and only if the number of votes is greater than a threshold value t. The DPoS mechanism consists of three main parts. 1) node selection, 2) block generation, and 3) block verification. The three components are described in detail.

1) **Node Selection**: the n manager nodes are selected based on the reputation of educational credentials in Section II-B3. The selected manager nodes are responsible for maintaining the blockchain and validating the blockchain generation. Each manager node has equal weights in the consensus mechanism.
2) **Block Generation**: the only manager nodes are responsible for validating the received transactions and packing them into the blockchain. Other committee nodes monitor the generated transaction in the blockchain network. Each manager node takes turns acting as a proposer to generate a block, and other manager nodes make voting mechanisms on the blocks generated by the proposer to ensure that most of the manager nodes agree on the block generation.
3) **Block Verification**: the manager nodes verify the generated block and broadcast the verification result *True* to other manager nodes. If the block is verified successfully. When a manager node receives a result *True* from managers nodes more than a threshold value. The block is verified successfully, and the current block is added to the blockchain.

IV. Experiment and Performance Evaluation

The system parameter is $\lambda = 256$. The experiment uses Geth[1] to build a private blockchain on a Ubuntu system. The smart contract is realized by the Solidity language and uploaded to a private Ethereum blockchain. The Web3.py[2] library is used to interact with the smart contract on the blockchain to test the time cost of sending transactions. We implemented the proxy re-encryption algorithm and the searchable encryption in Python.

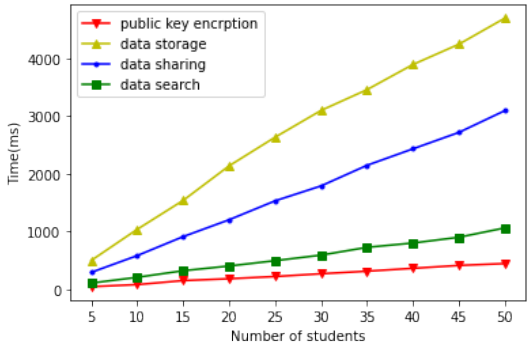

Fig. 2. Comparison of computational cost.

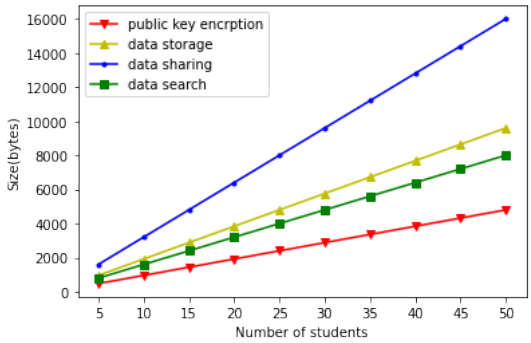

Fig. 3. Comparison of communication cost.

Since the computational cost is related to the number of students' credentials sn, the experiment evaluates these algorithms by setting $sn \in \{5, 10, ..., 50\}$. Fig 2. shows the computational cost of the data storage, data search, and data sharing method. The results show that the computational cost increases linearly with the increase in the number of students' credentials. Compared to using the public key encryption algorithm [7, 10], this proxy re-encryption algorithm contains a higher computational cost with an average increment of 72% in data storage and 53% in data sharing, but improves the security for educational credential sharing and solves the problem of non-verifiability of the re-encryption ciphertext. 3. shows the communication cost of the scheme increases linearly with the number of students; however, this scheme has more communication cost than the public key encryption, but this scheme improves the security of educational credential sharing.

[1]https://github.com/ethereum/go-ethereum

[2]https://github.com/ethereum/web3.py

REFERENCES

[1] L. J. Børresen, E. Meier, and S. A. Skjerven, "Detecting fake university degrees in a digital world," in *Corruption in Higher Education*. Brill, 2020, pp. 102–107.

[2] S. E. Eaton and J. J. Carmichael, "Fake degrees and credential fraud, contract cheating, and paper mills: Overview and historical perspectives," *Fake Degrees and Fraudulent Credentials in Higher Education*, pp. 1–22, 2023.

[3] M. Turkanović, M. Hölbl, K. Košič, M. Heričko, and A. Kamišalić, "Eductx: A blockchain-based higher education credit platform," *IEEE Access*, vol. 6, pp. 5112–5127, 2018.

[4] M. M. Arcinas, "A blockchain based framework for securing students' educational data," *Linguistica Antverpiensia*, vol. 2021, no. 2, pp. 4475–4484, 2021.

[5] W. Gräther, S. Kolvenbach, R. Ruland, J. Schütte, C. Torres, and F. Wendland, "Blockchain for education: lifelong learning passport," in *Proceedings of 1st ERCIM Blockchain workshop 2018*. European Society for Socially Embedded Technologies (EUSSET), 2018.

[6] A. Ayub Khan, A. A. Laghari, A. A. Shaikh, S. Bourouis, A. M. Mamlouk, and H. Alshazly, "Educational blockchain: A secure degree attestation and verification traceability architecture for higher education commission," *Applied Sciences*, vol. 11, no. 22, p. 10917, 2021.

[7] H. Li and D. Han, "Edurss: A blockchain-based educational records secure storage and sharing scheme," *IEEE Access*, vol. 7, pp. 179 273–179 289, 2019.

[8] L. Liu, M. Han, Y. Zhou, R. M. Parizi, and M. Korayem, "Blockchain-based certification for education, employment, and skill with incentive mechanism," *Blockchain cybersecurity, trust and privacy*, pp. 269–290, 2020.

[9] Z. Li and Z. Ma, "A blockchain-based credible and secure education experience data management scheme supporting for searchable encryption," *China Communications*, vol. 18, no. 6, pp. 172–183, 2021.

[10] R. A. Mishra, A. Kalla, A. Braeken, and M. Liyanage, "Privacy protected blockchain based architecture and implementation for sharing of students' credentials," *Information Processing & Management*, vol. 58, no. 3, p. 102512, 2021.

[11] D. Hou, J. Ma, Z. Peng, J. Zhang, and X. Zhu, "Privacy-preserving educational credentials sharing based on blockchain and proxy re-encryption," in *Computer Science and Education*, W. Hong and Y. Weng, Eds. Singapore: Springer Nature Singapore, 2023, pp. 119–130.

[12] F. Schuh and D. Larimer, "Bitshares 2.0: general overview," *accessed June-2017.[Online]. Available: http://docs. bitshares. org/downloads/bitshares-general. pdf*, 2017.