

Personal data sharing model with legal validity in blockchain-based secure environment

Su Jin Shin*, Sang Uk Shin**

*Department of Information Security, Graduate School, Pukyong National University, Republic of Korea

**Division of Computer Engineering and Artificial Intelligence, Pukyong National University, Republic of Korea
 inuin9014@pukyong.ac.kr, shinsu@pknu.ac.kr

Abstract— This paper proposes a data sharing model that ensures the legal validity of data provided by individuals as legal evidence within a blockchain environment. The proposed model applies Distributed Pseudo-Random Functions (DPRFs) to prevent all participants, except the data requester, from obtaining the original data. Based on these characteristics, the model guarantees integrity, confidentiality, and non-repudiation, thereby ensuring that data provided by individuals can possess legal validity as evidence.

Keywords— Blockchain, Data sharing, MPC, Legal evidence, Data security

I. INTRODUCTION

The recognition of any data as valid evidence in court signifies that the data possesses admissibility and sufficient weight [1]. In other words, for data to be accepted as legal evidence, its authenticity, integrity, identity with the original, and reliability must be assured. The data presented can include public information from the prosecution, police, and courts, as well as data provided by individuals. For instance, in the case of a vehicle accident, acceptable evidence may include road CCTV footage provided by the police, as well as dashcam recordings from a witness's vehicle. Research on data sharing models that ensure the effectiveness of legal evidence in a secure environment has been conducted for a long time. Existing studies have been designed based on blockchain environments to ensure a secure setting. However, these studies have focused on protecting the identity or authority of users seeking access to the data rather than protecting the data itself [2],[3]. There is a possibility that shared data may be exposed or manipulated. In other words, the properties of legal evidence, such as integrity and identity with the original, may be violated, which reduces the likelihood of being accepted as legal evidence. Therefore, this paper proposes a model that focuses on protecting original data so that data provided by individuals in a secure environment based on blockchain can be adopted as legal evidence. The proposed model performs encryption and decryption using values generated by participants involved in the computation process, ensuring that only the data requester can obtain the original data. Additionally, to prevent malicious

attackers from colluding with participants in the computation process in advance, the proposed scheme employs a random leader selection technique, such as that used in Algorand [4], in order to select random computation nodes. Finally, to encourage honest behavior of participants in the sharing model, a reward mechanism is employed, and requests for dispute resolution from participants are addressed through smart contracts, which identify dishonest participants and impose penalties on them. In this way, the proposed model ensures the resistance to collusion attack, fairness, and the effectiveness of legal evidence.

This paper is organized as follows: Chapter 2 outlines the cryptographic techniques employed in the proposed model, Chapter 3 details the components and operational steps of the proposed model, Chapter 4 provides a security analysis of the proposed model, and Chapter 5 concludes the paper.

II. CRYPTOGRAPHIC PRIMITIVE

This paper presents a protocol based on Distributed Pseudo-Random Functions (DPRFs), a type of Multi-party Computation (MPC). DPRFs are a cryptographic technique that enables computing participants to collaboratively compute a Pseudo-Random Function (PRF) [5]. They guarantee pseudo-randomness, ensuring that even if a malicious participant compromises other participants, the evaluation results remain confidential and are disclosed only to the evaluator. Furthermore, DPRFs ensure that the evaluation is independent of the participants, thereby maintaining consistency. The DPRF framework consists of three algorithms that satisfy the consistency property:

- $Setup(1^k, n, t) \rightarrow ((sk_1, \dots, sk_n), pp)$: The algorithm generates n secret keys $sk_1, \dots, sk_n$ and public parameters pp .
- $Eval(sk_i, x, pp) \rightarrow z_i$: The algorithm generates pseudo-random shares for the given values.
- $Combine((i, z_i)_{i \in S}, pp) =: z/\perp$: The algorithm combines the partial shares $\{z_i\}_{i \in S}$ from the entities in the set S to generate z .

The algorithm involves key shares of the PRF that each entity possesses, and they collaborate to compute the function [5]. However, each entity does not know the original key. The

proposed model uses three algorithms based on DPRFs: $(n, t) - DPRF$ protocol $DP := (DP.Setup, Eval, Combine)$, a cryptographically secure PRG (pseudo-random generator), and a *Commitment scheme*: $\Sigma := (\Sigma.Setup, Com)$ that possesses the properties of hiding and binding.

III. PROPOSED MODEL

This paper proposes a model for sharing legal evidence within a secure environment based on a consortium blockchain. The participants in the proposed model include evidence data owners, data requesters who seek to share evidence data necessary for resolving cases, and entities involved in the computation process. The data requesters and computation participants may consist of institutions related to the case, such as prosecutors, police, courts, and government agencies. Additionally, to enhance the efficiency of transaction data management, external storage solutions like Inter Planetary File System (IPFS) are employed. This section outlines the main components of the proposed model and explains the operational process step by step.

A. System Architecture

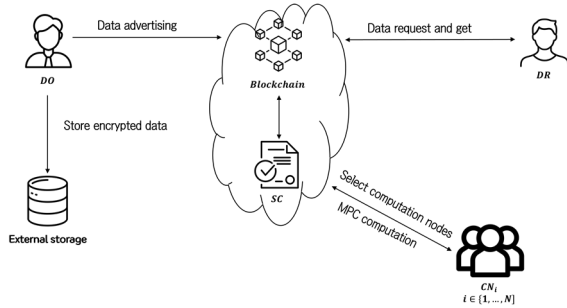


Figure 1. An overview of the proposed model

Figure 1 shows an overview of the proposed model. The main components of the proposed model are as follows:

- **Data Owner (DO):** The DO is the entity that deploys smart contract on the blockchain network and is a participant who wants to sell the data they have collected. The DO is an entity that performs encryption and wishes to receive a fair amount of payment for the data they provide.
- **Data Requester (DR):** The DR is a participant who wants to purchase data and engages in the protocol by distributing advertisements on the blockchain network. The DR wants to receive the correct data that corresponds to the amount they have paid.
- **Computation Nodes (CN):** The CN is a node that participates in the computation process among the nodes that are part of the blockchain network, randomly selected by a smart contract. It generates values used for encryption/decryption, and if the sharing process is correctly completed, it receives incentives from the DO and the DR.
- **Smart Contract (SC):** The SC provides automation and programming capabilities, which encapsulate the complex operations of nodes in a distributed blockchain system. In the proposed model, the SC randomly selects

computing nodes from the participating nodes in the blockchain network. Additionally, when receiving dispute resolution requests from participants, it identifies dishonest participants and carries out appropriate penalties.

B. Assumptions and Threat Model

The proposed model facilitates the sharing of personal data with legal validity within a consortium blockchain environment. Therefore, it inherits the security features of the blockchain platform, including integrity and traceability, as well as the use of smart contracts. Additionally, it is assumed that the primary components of the proposed model do not trust one another, prioritize their own interests, and act rationally. This means that they may engage in deceptive actions for their own benefit. The entity responsible for performing the actual computation process in the proposed model is not fully trusted. Therefore, it is assumed that these entities are selected using random leader selection techniques, such as those employed by Algorand [4], which provide a level of security. To implement these techniques, it is assumed that there is a sufficient number of nodes willing to participate in the computation process on the blockchain network. The attacker is presumed to be capable of causing static corruption [6], which can disrupt entities only before the protocol execution, and only considers the stand-alone model [7]. In other words, the attacker can compromise protocol participants solely prior to the execution of the protocol.

C. Details of the Proposed Model

The proposed model is designed based on a consortium blockchain environment and can be divided into the preprocessing and shared data advertising stage and the data sharing stage. The SC performs the dispute resolution stage when it receives a dispute resolution request from each participant. Figure 2 below shows the process flow of the proposed model.

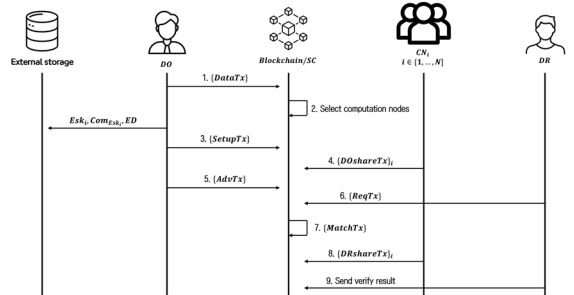


Figure 2. The processing flow of the proposed model

1) Preprocessing and Shared Data Advertising: The preprocessing and shared data advertising stage proceeds in the following order: the data disclosure stage, the random selection stage of the computation nodes, the interaction stage between DO and CNs, and the encryption and advertising stage.

- **Data disclosure step:** The DO generates the secret key sk_i , $i \in \{1, \dots, N\}$ and public parameters $pp := (pp_{DP}, pp_{COM})$ using the $DP.Setup$ algorithm and

Σ . *Setup* algorithm, where N denotes the number of computation nodes to be selected. Then, the DO generates the commit value Com_D for the original data D using the Com algorithm. The DO distributes the $\{DataTx\}$ transaction through the SC , which includes metadata about D , Com_D , N , and values needed for selecting computation nodes.

- Random selection step of the computation nodes: Among the participating nodes in the blockchain network, nodes that wish to participate in the computation process send a transaction $\{SignTx\}_j$ to the SC , which includes their signature $Sig_j(\{DataTx\})$. The SC uses a random computation node selection method, similar to the random leader selection technique of Algorand [4], to select N nodes as computation nodes. The SC updates the information of the selected nodes in the computation node list L and distributes it to the blockchain network.
- Interaction step between the DO and the CN_i s: The DO generates $Esk_i \leftarrow PK.Enc(sk_i, PK_{CN_i})$, $i \in \{1, \dots, N\}$ and then generates the commit value Com_{Esk_i} for Esk_i . The DO deploys a $\{SetupTx\}$ transaction through the SC , which includes the location URI_{Esk} where (Esk_i, Com_{Esk_i}) pairs are stored in external storage. After the CN_i obtains its Esk_i through the URI_{Esk} of $\{SetupTx\}$, it verifies the commit value for Esk_i . If the verification fails, the CN_i requests dispute resolution from the SC . If the verification is successful, the CN_i obtains $sk_i \leftarrow PK.Dec(Esk_i, SK_{CN_i})$ and then computes the share value used for encryption as $z_i \leftarrow Eval(sk_i, DO || Com_D, pp)$. The CN_i performs $Ez_i \leftarrow PK.Enc(z_i, PK_{DO})$ and sends $\{DOshareTx\}_i$, which includes the signature Sig_{Ez_i} to the SC .
- Encryption and advertising step: The DO obtains Ez_i through $\{DOshareTx\}_i$ and verifies the signature Sig_{Ez_i} . If the verification result fails, the DO requests dispute resolution from the SC . If the verification result is successful, the DO computes $z_i \leftarrow PK.Dec(Ez_i, SK_{DO})$. Then, the DO generates $w \leftarrow Combine(\{(i, z_i)\}_{i \in S}, pp)$ and creates the encrypted original data $ED \leftarrow PRG(w) \oplus (D || \rho)$. Based on this, the DO generates the ciphertext $C := (DO, Com_D, ED)$ and the signature Sig_C . The DO distributes a transaction $\{AdvTx\}$ containing the external storage location URI_C where C is stored and Sig_C through the SC .

2) Data Sharing: The data sharing step proceeds in the following order: Parsing step, interaction stage between DR and CN_i s, and finally the decryption and reward and punishment step.

- Parsing step: The DR selects one of the distributed $\{AdvTx\}$ and verifies the signature Sig_C included in that transaction. If the verification fails, the DR requests dispute resolution from the SC . If the verification is successful, the DR parses $C := (DO, Com_D, ED)$ and generates $x := (DO || Com_D)$. Then, the DR generates

the commit value Com_x for x and sends a $\{ReqTx\}$ transaction containing this to the SC . The SC checks the validity of the transaction and account information, and then distributes $\{MatchTx\}$ to indicate that the DR and the DO have been matched.

- Interaction Step between the DR and the CS s: The CN_i generates a commit value Com'_x for the x^* it holds and then verifies it. If the verification fails, the CN_i sends a dispute resolution request to the SC . If the verification is successful, the CN_i generates $z_i \leftarrow Eval(sk_i, DO || Com_D, pp)$ and generates $Ez_i \leftarrow PK.Enc(z_i, PK_{DR})$. The CN_i generates a signature Sig_{Ez_i} for Ez_i and distributes $\{DRshareTx\}_i$ through the SC .
- Decryption and reward and punishment step: The DR performs verification on Sig_{Ez_i} obtained through $\{DRshareTx\}_i$. If the verification fails, the DR sends a dispute resolution request to the SC . If the verification is successful, the DR generates $z_i \leftarrow PK.Dec(Ez_i, SK_{DR})$ and computes $w \leftarrow Combine(\{(i, z_i)\}_{i \in S}, pp)$. Then, the DR obtains $(D || \rho) \leftarrow PRG(w) \oplus ED$ and generates a commit value Com'_D for verification. If the verification fails, the DR sends a dispute resolution request to the SC . If the verification is successful, the SC distributes appropriate rewards to each participant and completes the sharing process.

3) Dispute Resolution Step: The dispute resolution stage is carried out only when the SC receives a dispute resolution request from each participant. The SC identifies dishonest participants and imposes appropriate penalties and compensations accordingly.

- In the case of receiving a dispute resolution request from CN_i : The SC verifies the validity of the transaction through the timestamp of the received transaction. If it is invalid, the transaction and the party that sent the transaction are excluded from the protocol. After that, the SC performs a verification of the commit value. If the value generated by the SC matches the value generated by CN_i , the SC determines that another participant is dishonest and imposes a penalty by confiscating that participant's deposit and distributing it to other honest participants, thus concluding the transaction. If the value generated by the SC does not match the value generated by CN_i , a penalty is imposed on CN_i , and the transaction is terminated only if the number of dishonest CN_i exceeds a certain threshold.
- In the case of receiving a dispute resolution request from the DO : The SC verifies the validity of the transaction through the timestamp of the received transaction. If it is invalid, the transaction is excluded from the protocol. After that, the SC performs a verification of the signature value. If the verification fails, the SC determines that the DO is a dishonest participant and imposes a penalty on the DO before terminating the transaction. If the verification is successful, the SC

determines that CN_i is a dishonest participant and imposes a penalty. If the number of dishonest CN_i exceeds the threshold, the transaction is terminated.

- In the case of receiving a dispute resolution request from DR : The SC verifies the validity of the transaction through the timestamp of the received transaction. After that, the SC performs a verification of the signature value. If the verification fails, the SC determines that DR is a dishonest participant and imposes a penalty on DR , terminating the transaction. If the verification is successful, the SC determines that CN_i is a dishonest participant and imposes a penalty. If the number of dishonest CN_i exceeds the threshold, the transaction is terminated.
- In decryption and reward and punishment step, when a dispute resolution request is received from DR : The SC verifies the validity of the transaction through the timestamp of the received transaction. After that, the SC informs that a dispute has occurred regarding D through the transaction. The DO , who confirmed the transaction, and CN_i disclose D and z_i , respectively, and the SC uses this information to identify the dishonest participant. The SC then carries out penalties against the identified dishonest participant and terminates the transaction.

IV. SECURITY ANALYSIS

In this chapter, we perform a security analysis of the proposed model, examining it from the perspectives of collusion attacks, fairness, and its effectiveness as legal evidence.

A. Resistance to Collusion Attack

An attacker can engage in malicious activities by colluding with a computing node. To execute this, the attacker must identify a node to collaborate with in malicious activities before the protocol begins, and the chosen node must be designated as the computing node. However, the proposed model is built upon the use of cryptographically secure hash functions and signature techniques, as well as the security properties of Algorand [4]. Consequently, it is difficult to predict in advance which nodes participating in the blockchain network will be selected as computing nodes. As a result, the likelihood of success of a collusion attack in the proposed model is negligible.

B. Fairness

In the proposed model, when the SC receives a dispute resolution request from a transaction participant, the SC performs the dispute resolution process through the verification of commitments or signatures. Through this, the SC detects dishonest transaction participants and imposes penalties by confiscating the deposits of those dishonest participants and distributing them to honest transaction participants. The proposed model is based on cryptographically secure hash functions, commitment schemes, and signature schemes, ensuring the security of the cryptographic mechanisms and the security of smart contracts on the blockchain. Based on these

characteristics, the proposed scheme ensures fairness by imposing appropriate penalties or rewards on participants DO , DR , and CN_i .

C. The Effectiveness as Legal Evidence

The proposed model generates and transmits values used for encryption/decryption or signatures of encrypted values, or commit values, in a blockchain environment. Additionally, the entire processes of data transfer and interaction between participating entities are recorded as transactions. In the encryption/decryption process of the proposed model, the value z_i or w generated by CN_i is utilized. If CN s above the threshold collude, they could obtain the shared data through malicious behavior; however, the proposed model guarantees protection against pre-collusion attacks. Therefore, only the DR can obtain the original shared data. In this way, the proposed model ensures characteristics such as integrity, non-repudiation, confidentiality, and traceability, thereby guaranteeing that the shared data within the proposed model serves as legal evidence.

V. CONCLUSIONS

This paper proposes a model for securely sharing evidence while maintaining its legal validity within a blockchain environment. The proposed model enhances data encryption by applying DPRFs to perform encryption and decryption based on values generated by computing nodes. Additionally, to mitigate data exposure risks associated with collusion attacks from malicious attackers, a random leader selection technique is employed to thwart premeditated collusion efforts. All values transmitted during the sharing process of the proposed model are recorded as transactions, and these transmitted values also include digital signatures. This approach ensures integrity, non-repudiation, and traceability. By guaranteeing these characteristics, the data shared within the proposed model maintains its validity as legal evidence.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No.2022R111A3063257), and supported as a 'Technology Commercialization Collaboration Platform Construction' project of the INNOPOLIS FOUNDATION (Project Number: 1711202494).

REFERENCES

- [1] S.J Lee, Introduction to Digital Forensics, Revised edition., Irun Publishing, 2015
- [2] W. Du, H. Liu, G. Luo, J. Zhang and W. Xu, "A Consortium Blockchain-Enabled Evidence Sharing System for Public Interest Litigation", *Journal of Global Information Management (JGIM)*, vol.31, pp. 1–19, 2023.
- [3] M. Li, C. Lal, M. Conti and D. Hu, "LEChain: A blockchain-based lawful evidence management scheme for digital forensics", *Future Generation Computer Systems*, vol. 115, pp. 406–420, 2021.
- [4] J. Chen and S. Micali, "Algorand", *arXiv preprint arXiv:1607.01341*, 2016.
- [5] S. Agrawal, P. Mohassel, P. Mukherjee and P. Rindal, "DiSE: distributed symmetric-key encryption", *Proceedings of the 2018 ACM*

SIGSAC conference on computer and communications security, pp. 1993–2010, 2018.

- [6] M. Abe and M. Ohkubo, “A framework for universally composable non-committing blind signatures”, *International Journal of Applied Cryptography*, vol. 2, pp. 229–249, 2012.
- [7] A. Y. Lindell, “Adaptively secure two-party computation with erasures”, *Cryptographers’ Track at the RSA Conference*, pp. 117–132, 2009.



Su Jin Shin received her B.E. degree in Dept. of Software and Artificial Intelligence from Pukyong National University, Republic of Korea in 2023. She is currently a master course student in Department of Information Security, Graduate School from Pukyong National University. Her research interests are related with blockchain security, applied cryptography, and vehicle security.



Prof. Sang Uk Shin received his M.S. and Ph.D. degrees from Pukyong National University, Busan, Korea in 1997 and 2000, respectively. He worked as a senior researcher in Electronics and Telecommunications Research Institute, Daejeon Korea from 2000 to 2003. He is currently a professor in Division of Computer Engineering and Artificial Intelligence, College of Information Technology and Convergence, Pukyong National University. His research interests include cryptographic protocol, blockchain security, mobile and wireless network security, and digital forensics.