Selective Layered Blockchain Framework for Privacy-preserving Data Management in Low-latency Mobile Networks

*Sun-Woo Yun, Eun-Young Lee, Il-Gu Lee**

Department of Future Convergence Technology Engineering, Sungshin Women's University, The Republic of Korea
nus0205@naver.com, o.lee.eunyoung@gmail.com, iglee@sungshin.ac.kr

Abstract

With the gradual development of Fourth Industrial Revolution technologies, such as artificial intelligence, the Internet of Things, and big data, and the considerable amount of data in mobile networks, low-latency communication and security management are becoming crucial. Blockchain is a data-distributed processing technology that tracks data records to support secure electronic money transactions and data security management in a peer-to-peer environment without the need of a central trusted authority. The data uploaded to the blockchain-shared ledger are immutable, making tracking integrity preservation facile. However, blockchain technology is limited because it is challenging to utilize in the industry owing to its inability to correct data, even when inaccurate data are uploaded. Accordingly, research on blockchain mechanisms that consider privacy-preserving data management is required to commercialize blockchain technology. Previously, off-chain, blacklist, and hard-fork methods have been proposed; however, their application is challenging or impractical. Therefore, to protect privacy, we propose a layered blockchain mechanism that can correct data by adding a buffer blockchain. We evaluated the latency, security, and space complexity of layered blockchains. The security and security-to-latency ratio for data management of the selective layered blockchain is 2.2 and 11.3 times higher than the conventional blockchains, respectively. The proposed selective layered blockchain is expected to promote the commercialization of blockchain technologies in various industries by protecting user privacy.

Keywords: Layered blockchain, Privacy, Security, Data management, Data correction

1 Introduction

In 2008, Satoshi introduced blockchain technology to prevent double spending through peer-to-peer networks. Satoshi's paper reasoned that blockchains could facilitate online financial transactions between stakeholders without the intervention of a central trust institution that could be superseded by the cryptography and consensus technology of blockchains [1]. Therefore, as a peer-to-peer distributed computing technology, the blockchain solves the common Byzantine problem encountered in unreliable networks.

Blockchain technology is beneficial in preventing data falsification and reducing transaction costs. Blockchains promote economic value creation and improve security and efficiency in various sectors such as finance, logistics, distribution, healthcare and energy [2]. Blockchain is a fundamental technology in the Fourth Industrial Revolution because it can safely store data without relying on trusted third parties [3]. However, for this technology to be commercialized as a core infrastructure technology, we must address its limited scalability and security. Although the inability to modify or delete data already stored in a blockchain is an excellent feature in terms of reliability, it is also an obstacle to the secure utilization and expansion of blockchain technology in various industries. In industries that store and utilize sensitive personal information in considerable quantities, such as the Internet of Things and mobile networks, blockchain can prevent the loss of personal data [4]. However, the immutability and integrity of blockchain technology disregard the right-to-be-forgotten principle, and transparency can trace the transaction data of the nodes. Malicious attackers can trace the flow of transactions and reveal the true identities of users based on data mining. In particular, uploading sensitive data to a blockchain may conflict with privacy laws such as the General Data Protection Regulation (GDPR) [5], thus limiting the active use and commercialization of blockchain as a core infrastructure technology [6]. Accordingly, more research is required on the feasibility and application methods when applying blockchain technology in industries that utilize sensitive personal information, such as health, medicine, insurance, politics, mobile and finance [7]. Recently, several solutions have been proposed to protect blockchain privacy, including those based on the off-chain, blacklist, and hard-fork methods [3]. Because the off-chain method stores personal information outside a block, it is not decentralized and vulnerable to hacking and maintenance. The blacklist method uses an encrypted key to access personal information such that data are inaccessible without the key; however, records can still be traced [8]. Finally, the hard-fork method authorizes the removal of personal information by dividing it into an old blockchain and a new version; however, it is impractical as a commercial technology because it is difficult to implement and has structural limitations by changing the existing conventional system [9]. Consequently, research into blockchain technology that can protect privacy while overcoming the limitations of existing solutions is necessary.

*Corresponding Author: Il-Gu Lee; E-mail: iglee@sungshin.ac.kr
DOI: 10.53106/160792642023072404006

This study proposes a layered blockchain architecture composed of multiple blockchains. In a layered blockchain architecture, layers operate as buffers to ensure that transactions are not approved immediately. Transaction data can be rectified until they are uploaded to the top layer of the blockchain. In a layered blockchain, the number of layers can be adjusted flexibly according to the application field. As the number of layers increases, more data are rectified to prevent privacy leakage. Despite the numerous studies on blockchain architectures composed of several layers, few studies have been conducted on layered blockchains with a specific focus on privacy protection. Most studies on layered blockchains involve providing permission based on members' roles and proposing measures to improve the performance or efficiency of the mechanism rather than improving privacy. In contrast, studies have been conducted on layered blockchains to ensure privacy and the right to forgetting. Blockchain-based services cannot modify or delete contents and personal information that have already been uploaded owing to their history of data deletions and updates, and these features disregard the GDPR. Therefore, a three-layered blockchain architecture was proposed for blockchain services that provide insertion and deletion functions without compromising the decentralization or integrity of the blockchain [8]. The architecture comprises three layers: service, link, and content data storage. This study resolves the difficulty of index modification by introducing additional blockchains to manage the connections between the content and content indices. The architecture is also expected to provide blockchain services that can be used for personal content management while enabling data to be traced, modified, and deleted. However, the study [8] only noted various cases for the proposed architecture application based on the use case and failed to verify the performance improvement, security effectiveness, and system's implementation potential compared to conventional services. Therefore, a blockchain architecture that can be verified to protect privacy while maintaining the advantages of a layered blockchain when implemented based on objective evaluation indicators is necessary.

Therefore, this study proposes a layered blockchain architecture that can prevent privacy leakage in uploaded data for a certain period by adding a blockchain for buffers. Furthermore, we conducted verification and comparative analysis studies on privacy protection capabilities and system performance in conventional and off-chain blockchains.

The main contributions of this study are as follows:

- A layered blockchain architecture with an additional blockchain for buffers is proposed to protect the privacy and the right to be forgotten on mobile networks while preserving the immutability and integrity of the blockchain.
- The proposed layered blockchain is simulated and its effectiveness evaluated in terms of latency, security, and space complexity.
- A selective layered blockchain methodology that can overcome the complexity limitations of layered blockchain structures is proposed.

The remainder of this study is organized as follows. In Section 2, we compare and analyze previous studies on blockchain for privacy. Section 3 introduces a layered blockchain architecture with a data correction function and compares it with conventional blockchain and hyperlink blockchain methods. Section 4 presents the performance and security evaluation of a layered blockchain using a simulator. Finally, Section 5 presents the conclusions and a discussion of future research directions.

2 Background and Related Works

This section introduces conventional and hyperlink blockchain techniques and prior studies on the right to be forgotten and privacy are compared and analyzed.

2.1 Conventional Blockchain

Electronic assets in the digital environment can be freely duplicated, and cloned copies are indistinguishable from the originals. Therefore, if a currency is replicated and used indefinitely, it cannot fulfill its role as a currency. The duplication and reuse of the same currency are called double spending [10]. Previously, trusted intermediaries approved transactions with their trusted counterparts to prevent double spending. Bitcoin was the first cryptocurrency to use a blockchain structure, enabling it to be traded between individuals without a trusted third party in a digital environment [1]. The reliability of the blockchain system can be guaranteed based on the proof of work, which refers to discovering a specific value, where the resulting value of the hash function SHA256 begins with a certain number comprising zero bits. The difficulty of the proof-of-the-work process is determined based on the block generation time and is adjusted by changing the average target amount based on the hourly average number of blocks. Given that a block contains the previous block information, the computational time required for falsification increases exponentially with an increasing number of blocks within a blockchain. Therefore, a blockchain is secure if a group of honest nodes controls more CPU power than malicious nodes.

Although the irreversible nature of blockchain is essential for ensuring reliable transactions between individuals in a digital environment, its technical properties may result in privacy issues [11]. In particular, managing personal or sensitive information is challenging because it cannot be altered or changed after being uploaded to a blockchain [12]. Therefore, a method that guarantees the right to be forgotten is required for the blockchain technology to be more suitable for practical applications. Numerous studies have been conducted to ensure the right-to-be-forgotten in conventional blockchains. A multilayered blockchain framework consisting of six layers enables users to set the scope of data disclosure, duration of sharing, and redistribution criteria [13]. This framework can also protect users' personal information using k-anonymity and differential privacy. The temporal rolling blockchain method proposes a feature for blockchain nodes to store data for a predetermined period, subsequently deleting old data [14]. The entire ledger does not require storage by full nodes and is maintained by all nodes, thus reducing the storage burden while protecting privacy to some capacity. However, these methods cannot correct the data stored in the blockchain, perform data correction, and provide

only an auxiliary means of protecting personal information. Specifically, the real-life application of blockchain remains limited because it disregards the GDPR, which requires users to be able to modify and delete personal data.

2.2 Hyperlink Blockchain

Blockchain was developed to enable various data to be inputted based on the purpose of the data and the designer's intentions. Numerous studies have been conducted to develop new models using the blockchain (on-chain) and databases (off-chain). For example, a two-layered system was developed to verify data integrity by operating a blockchain and an off-chain database in parallel and using one of those layers to provide access control [15]. However, although the data integrity was verified by combining the database and blockchain, this method cannot support modification or deletion of data inputted into the blockchain.

The ultimate objective of blockchain is to share reliable information [16]. In this study, we selected a hyperlink-based on–off blockchain model as the primary solution for the right to be forgotten and privacy; implementing and supporting modification and deletion functions directly is feasible. The proposed model was analyzed and compared with a conventional blockchain. Figure 1 depicts a schematic diagram of the hyperlink method.

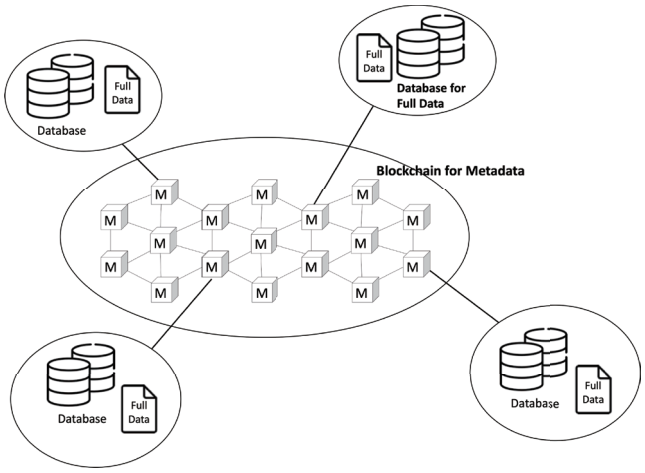


Figure 1. Hyperlink method

The hyperlink method manages data using an existing centralized database by inserting metadata into a blockchain and linking the two through hyperlinks [17]. Specifically, managing data in a centralized database makes modifying them more manageable; moreover, a blockchain that shares only simple data, such as keywords and abstracts is utilized. A blockchain that permanently stores data can prevent privacy and information leakage by inputting only the minimum amount of information necessary for searching.

A centralized database supports all creation, reading, update, and delete functions, so it can manage significant data systematically and conveniently. Conversely, because the system is centralized, it is prone to single point of failure, wherein if a part of a system fails, it will stop the entire system from working [18]. The notable difference between a blockchain and a database is whether modification or deletion functions exist. The hyperlink method is an advanced method that combines the features of both blockchain and database techniques, where the metadata is managed in a blockchain layer and all data are managed with off-chain techniques based on utilizing hyperlinks. The modification and deletion functions of the database ensure the right to information so that the subjects can control their own data, whereas the blockchain provides safe administrative management. However, this hyperlink method has limitations for commercialization owing to legal conflicts from a privacy perspective. In a blockchain model that manages personal information, sensitive information is collected from an existing off-chain database and the corresponding indirect information is shared through the blockchain. A database that manages information and a blockchain that shares indirect information cannot be viewed as equivalent system layers, which is an act of entrustment. Therefore, the legal relationship between main-chain and off-chain data is ambiguous and may hinder the development of this method.

2.3 Privacy and Data Correction

In this section, research on blockchain techniques for the right to be forgotten and privacy are compared and investigated, and current solutions and their limitations are analyzed. Previously, blockchain for privacy protection was described by focusing on deletion function; however, in this study, we propose a layered blockchain, including a data correction function to correct inappropriate privacy data. Table 1 summarizes the absence of data deletion and correction functions; key technologies; and considerations for prior research.

A previous study [8] proposed a multilayered blockchain architecture. The concept of a proposed layered blockchain was used as a technique for the right to be forgotten. Conversely, this study addresses the inability to correct uploaded content and personal information to delete and update the data history. The blockchain consists of three layers and introduces additional blockchains making managing the link between the content and a content index easier. Each of the three layers is responsible for storing the information used in content services, the information for the link between services and encrypted content, and the encrypted content and content information. As personal information and content can be modified and deleted, data correction is possible for data with privacy leakage; however, it may be utilized for other purposes and may present a risk for data manipulation. Among conventional blockchain techniques for the right to be forgotten and privacy, mechanisms that enable the deletion of only data that satisfy specific conditions have also been proposed. For example, data can be considered a living organism that ages over time to determine whether to preserve or delete data based on its usage history and rate [9]. However, this method may disregard invariance and integrity, which are unique features of the blockchain. Therefore, a mechanism for designing IT artifacts as prototypes and deleting old data has been proposed using the design science research approach while maintaining most of the primary features of blockchain technology [19]. Conversely, methods have been proposed to provide a distributed ledger that maintains invariance, integrity, and transparency; and protect privacy according

to the utilization of deletion functions using an existing centralized, off-chain database. For example, the scope is defined such that personally identifiable information (PII) and non-PII are stored in separate locations and executed within limits permitted by personal information regulations [20]. Similarly, methods for storing metadata in the main blockchain while processing sensitive data as reference have been studied [21]. A mutable blockchain that sets the mutability policy inside a smart contract can delete and modify blockchain data records by specifying active and non-active transactions in a series of transaction sets [22]. Users would not have access to non-active transactions, and can only refer to active transactions.

Thus, blockchain is highly likely to be used in various industries as a reliable data-sharing mechanism; however, technological improvements are still required for its commercialization. Although various studies are underway to satisfy GDPR privacy regulations and privacy requirements, additional research is required on mechanisms to ensure the performance and security of the blockchain architecture while preserving its integrity [23].

Table 1. Blockchain research on privacy and data correction

Reference	Year	Data deletion	Data Correction	Key technologies	Open issues
[8]	2020	O	O	Three-layered blockchain architecture consisting of public service blockchain, public service blockchain, and public contents blockchain for private contents management	Only use case-based concepts are presented without assessing implementation feasibility and performance
[13]	2019	X	X	A six-layered blockchain to secure the mobility data (Inspired by the Open System Interconnection model)	Only the user's data access can be controlled
[14]	2016	X	X	The feature used to reduce storage burden by deleting old data and storing only the latest data	Even if you delete the data of a blockchain node, it remains on another node
[15]	2015	Δ (only data from distributed hash tables and the centralized cloud can be deleted)	X	The off-blockchain key-value store to protect the privacy and control access to personal data	Only the user's data access can be controlled, and efficient data processing requires discussions
[17]	2019	Δ (only data from a centralized database can be deleted)	X	Management of blockchain-based subject data for storing and managing considerable data	Efficient data processing requires discussion and not decentralized
[19]	2019	O	X	Prototype information technology artifact to delete old data in the blockchain using the design science research approach	Only limited to old data without resolving privacy law violations
[20]	2018	O	X	Off-chain blockchain architecture using local databases and distributed ledger to separate storage space between PII (personally identifiable information) and non-PII	The life cycle of PII is maintained; however, personal information-related law is not entirely adhered to (i.e., users must be able to delete or correct their data)
[21]	2018	O	X	Modular blockchain architecture processing sensitive data and utilization of metadata	Vulnerable to hacking and maintenance, and not decentralized
[22]	2017	X	O	Setting mutability policy inside the smart contract, and replacing a vulnerability flow smart contract with a normal transaction	Right to be forgotten requires establishing before transaction distribution

3 Privacy-preserving Data Management on Layered Blockchain

This section describes the structure and principles of the proposed layered blockchain. Moreover, we compare the proposed mechanism with conventional blockchain and off-chain hyperlink methods.

3.1 Proposed Layered Blockchain Architecture and Privacy-preserving Data Management

Layered blockchain is a blockchain architecture composed of several layers. In this study, blockchain groups acting as buffers were appended systematically. The data of the lower blockchain layers are correctable. A layered blockchain can correct data until they are uploaded to the final blockchain layer, a public blockchain. The data correction period is *the block time* × (*the number of layers* – 1). The block time and number of layers can vary according to the blockchain platform, application area, and security level. For example, if the block time is 10 min and the number of layers equals three, the layered blockchain can correct the data for 20 min. In this study, we assumed a three-layer structure for the layered blockchain, as described below.

Data correction is a primary function for privacy protection that corrects inappropriate privacy data. Consequently, the privacy data includes sensitive data that should not be leaked, inappropriate data, incorrectly uploaded data, and malicious data. Layered blockchain performs numerous data correction as the number of layers. Among the uploaded data, we calculated privacy data based on the consensus among the nodes.

Concerning a layered architecture with a three-layered configuration, layers 1 and 2 consist of a private blockchain, and layer 3 consists of a public blockchain. It is assumed that layer 3 is transparent, and layers 1 and 2 maintain confidentiality. The composition of each layer can be defined based on the size of the blockchain network group, which expands as the stages progress. In layer 3, all members participate in the blockchain. In layers 1 and 2, all members are divided into several small blockchains. The number of nodes in each layer was assumed to be sufficiently large to guarantee a certain level of security. The layers in a layered blockchain can function complementarily to ensure integrity. The upper layer guarantees the security of the lower layer, and the lower layer can assess whether the integrity of the system is compromised through the upper layer. For example, if an attack occurs in layer 1 and illegally privacy data are uploaded, the original data and history of data corrections can be traced in the upper and lower layers. This aspect prevents privacy information from being uploaded to the upper layer and facilitates tracing specific data.

In a layered blockchain, the data correction function is defined as an "DC" function. Correcting data in transactions involves deleting and correcting existing data. The DC redistributes the contract for data determined to be incorrectly present based on an agreement between nodes and creates a redistribution history. Consequently, the latest contract can track the DC history by referring to the hash value of the previous contract. Data transactions before DC modify the contract's status to "Not uploaded" so that the inappropriate privacy data are not uploaded to the upper layer. The default value of the contract state is defined as "Uploaded," and altering the status from "Not uploaded" to "Uploaded" is not possible. If DC is not required, the status is maintained as "Uploaded," and data are uploaded to the upper layer. The DC functions can be executed from the moment the data are uploaded to each layer to the period when the blocks are created. The operation of the DC function of a layered blockchain is illustrated in Figure 2.

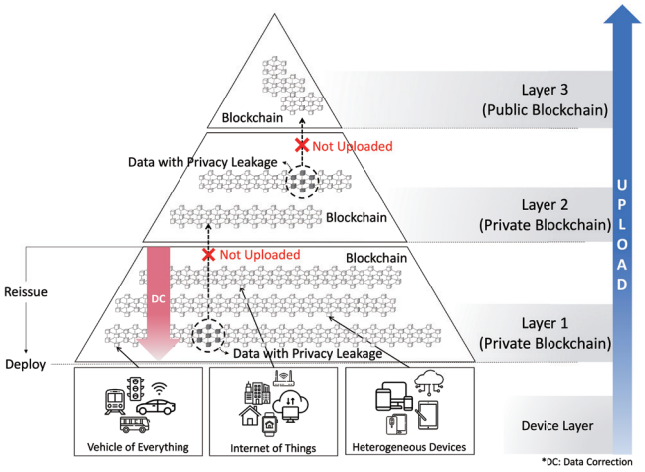


Figure 2. Layered blockchain architecture

When a user uploads new data, a contract is distributed to the private blockchain network in layer 1. The DC function of the layered blockchain can be performed until a block is created, and the privacy data transaction rate is calculated based on the agreement between nodes. Depending on the consensus algorithm, the data correction capability of each blockchain group may differ. When inappropriate privacy data are detected, an issue function is executed to redistribute the contract to the corrected data. Conversely, for data with privacy leakage before correction, the contract's status alters from "Uploaded" to "Not uploaded," and the data cannot enter the subsequent layer. After a certain period elapses to enable block creation, data with an "Uploaded" status (among the data in each layer) are uploaded to layer 2. The operation process of layer 2 is similar to that of layer 1. Only the deletion function can be executed in layer 2, which operates identically to layer 1 when requested. After a certain period elapses to create the block in layer 2, the data are uploaded to layer 3, corresponding to the final layer, after which it cannot be corrected. Specifically, the uploading to the final layer is similar to uploading data to a conventional blockchain; however, it includes a history of data corrections made in the lower layers that can be traced. In conclusion, the layered blockchain can lower the rate of privacy data uploaded to the public blockchain and minimize the scope of data disclosure to protect privacy and the right to be forgotten.

3.2 Comparison with Conventional Mechanisms

The transparency and integrity of the blockchain do not guarantee the right to be forgotten or the right of the information subject to control all related information.

Ensuring the privacy of uploaded data using conventional blockchain techniques is challenging, owing to their transparency and integrity. However, regarding a layered blockchain, it is possible to decrease the extent of privacy data leakage and improve security compared to conventional methods. This is because the data are uploaded to a public blockchain after they are filtered sufficiently in the private blockchain of the lower layers. The comparison between conventional, hyperlink, and layered blockchains are in Table 2.

Table 2. Comparison of the conventional, hyperlink, and layered blockchain

	Conventional blockchain	Hyperlink blockchain	Layered blockchain
Data deletion	Impossible	Disconnect link	Change contract status (uploaded → not uploaded)
Data correction	Impossible	DB modification → Upload metadata to the blockchain	Correct and delete the data
Management	Easy	Not easy	Easy
Period	None	Always	Predetermined period
Security	Data integrity ensured	Vulnerable database, single failure point	Data integrity ensured
Policy relevance	Users cannot modify and delete data when they desire Application is limited for sensitive information	The layer that stores data and the layer that stores metadata have a vertical structure in the form of entrustment Inadequate for use with personal and sensitive information	The multi-layered structure of equivalent layers, not as entrustment Application area expanded by providing additional time for modification and deletion

A conventional blockchain does not allow data to be deleted owing to its structural limitations. This method stores the original data in the blockchain, thus making it manageable and ensuring the integrity of the data and reliability of the distributed ledger. However, the inability of users to delete or correct information can result in legal conflicts. Furthermore, the limitation of the conventional blockchain is that it has constraints when applied in areas that concern sensitive data.

The hyperlink method stores original data in a database and uploads the metadata to the blockchain. Data correction is possible by modifying the original data stored in the database and reuploading the metadata to the blockchain. Data deletion is performed by deleting the original data stored in the database and terminating the link between the original data and the metadata. Although data can be corrected and deleted anytime, the integrity of the data is challenging to manage through this blockchain, and if the database is insecure, it can become a single point of failure. Moreover, entrusting data and data-link information to a non-equivalent layer is inappropriate for personal and sensitive information.

The layered blockchain proposed in this study considers data in private blockchains of lower layers based on a hierarchical structure, subsequently uploading the data to a public blockchain. Data corrections are performed by re-uploading the data to the lower layers, based on which the correction details can be traced. The inaccurate contract before DC alters the status to "Not uploaded" so that the data do not proceed to the upper layer. Original data are stored in a blockchain for easy management. Data corrections are performed for predetermined periods. Unlike in the entrustment approach, all layers are equivalent, and the application area of the blockchain is expanded by providing an additional period for data correction. Layered blockchains guarantee a higher degree of security than conventional techniques without disregarding the fundamental technical principles of blockchain technology.

Table 3 details the analyzed outcomes of the conventional blockchain, hyperlink method, and layered blockchain regarding general and sensitive data for the right to be forgotten.

Table 3. Analysis of the conventional, hyperlink, and layered blockchain methods for data

Conventional blockchain	Hyperlink blockchain		Layered blockchain
All data type	Non-sensitive data	Sensitive data	All data type
The right to be forgotten is not guaranteed	Metadata remain, and inference is possible but still difficult	Implementation is impossible owing to entrustment (or costs incurred for compensation)	Data remains and the right to be forgotten is guaranteed because of a narrow scope of disclosure

Conventional blockchain does not guarantee the right to be forgotten for either general or sensitive data. In the hyperlink method, general data remain in the metadata, and the original data may be inferred based on them, although this process remains challenging. Additionally, implementing the method with sensitive data is not feasible owing to legal problems originating from the vertical relationship between

the off-chain and blockchain. Conversely, the layered blockchain leaves data in the lower layers but still guarantees the right to be forgotten by restricting the scope of the disclosure. In addition, many studies have been conducted on utilizing blockchain technology, thus ensuring integrity to protect personal and sensitive information. For privacy protection, research has been conducted to improve the structure of blockchain systems, develop security methods for privacy estimation, and duplicate attacks or frameworks for application in specific industries. A dedicated privacy-preserving secured framework leverages blockchain and deep learning technologies to provide reliability two-level privacy, and intrusion detection modules [24]. The privacy-preserving threat intelligence framework combines blockchain and deep learning technologies [25]. Similar to these studies, several other studies have been conducted to prevent inferences or duplicate attacks by changing the data format to be protected by combining various deep learning modules and models with blockchain [24-25]. In blockchain-based privacy-preserving and private data-sharing schemes, data security, access control, and licensing functions can be provided based on the blockchain's smart contract to aid in sharing personal information as digital asset [26]. Therefore, studies on personal-information protection measures using blockchain have been conducted from various perspectives. However, there is a risk in storage consignment when using the off-chain method, and the data can be released to anyone. The layered blockchain proposed in this study does not define a response to a specific attack threat or data storage format as described above. However, sensitive information with minimal privacy leakage can be safely uploaded within the scope of the privacy policies. Contrarily, owing to the structural nature of a layered blockchain, the higher the number of transactions and layers, the greater is the overall complexity, which leads to performance degradation. Therefore, this study proposes a selective layered model among the methods to utilize the layered structure. This model can configure different numbers of layers, depending on the sensitivity of the data. Data sensitivity can be classified by a data management authority according to the application area and security level. Among the data classified according to security requirements, the more sensitive data passes through more layers. Selective layering provides scalability to flexibly utilize layered structures based on the nature of the data while improving technical performance.

4 Evaluation

We evaluate the performance, security, and space complexity of a layered blockchain using a simplified blockchain simulator in this section.

4.1 Evaluation Environment

We implemented the evaluation simulator using Python version 3.9.12. All simulations were performed on a MacOS 12.0.1 with 16 GB RAM and a 10-core CPU Apple M1 Pro chip.

For evaluation, we compared a layered blockchain with a conventional blockchain. The comparative models are conventional blockchain frameworks without data correction function [1, 12]. Data uploaded to a conventional blockchain cannot be modified because of their irreversibility. The proposed models are divided into layered and selective-layered blockchains. Both models are blockchains with multiple layers. In a layered blockchain, all data go through a set layer, and in a selective-layered blockchain, the number of layers the data go through varies regarding data sensitivity. As data sensitivity increases, the number of layers through which the data passes also increases.

According to the International Standardization/International Electrochemical Commission (ISO/IEC) 27001, companies should classify and manage their data [27]. Companies should establish data classification criteria based on application areas and security levels. Generally, the data are classified into three or four categories. The data classification criteria of selective layered blockchain refer to Federal Information Processing Standards 199 Confidentiality [28]. Low, moderate, and high levels are based on the potential impact of the data. In addition, we included a public level that was unaffected by leaks. Table 4 summarizes the data sensitivity levels defined in this study.

Table 4. Data sensitivity levels

Level	Definition	The number of layers which go through
Public	Data unaffected by leaks	One-layer
Low	Data expected to have a limited adverse effect	Two-layer
Moderate	Data expected to have a serious adverse effect	Three-layer
High	Data expected to have a serious or catastrophic adverse effect	Four-layer

There were four levels of data sensitivity. Data that can be revealed are at the public level and are passed through a single layer. Data to be protected, such as biometric and personal data, are classified into three levels depending on their impact. Data expected to have limited adverse effects are low and go through two stages. Data expected to have serious adverse effects were moderate and went through three stages. Data expected to have sever or catastrophic adverse effects were high and went through four layers.

In the top layer, all the members participate in the blockchain. The lower layers are divided into several small groups. For example, in layer 1 of the four-layered blockchain, eight groups process 12.5% of the total transaction issuances. In layer 2, the four groups process 25%. In layer 3, the two groups process 50%. In layer 4, one group processes 100%. Simultaneously as the number of layers is i, the transaction throughput by one group of layers i can be calculated using Equation 1:

$$Amount\ of\ transaction\ throughput = \frac{total\ transaction\ issuance}{2^{i-1}}. \quad (1)$$

Figure 3 depicts the operational flow and a block diagram of the evaluation model used in this study. Transactions are issued when users upload new data. Issued transactions create blocks that are uploaded to the chain. Each block contains an index, a timestamp, a preview hash, a hash, and data. A conventional blockchain cannot be modified even if there is a transaction with privacy leakage in the first uploaded data. A layered blockchain appends a process for selecting privacy leakage data based on node-to-node consensus. In the proposed model, the initial status of the data is set to "Upload." This status enables determining whether to redistribute or delete the block before uploading it to a higher layer. DC is performed differently depending on whether the data are corrected. If it is necessary to delete data, the status of these data is switched to "Not uploaded" and they are not uploaded to the next layer. If it is necessary to correct the data, the corrected data are redistributed to the subsequent layer. The data status before correction is switched to "Not uploaded" and they are not uploaded to the next layer.

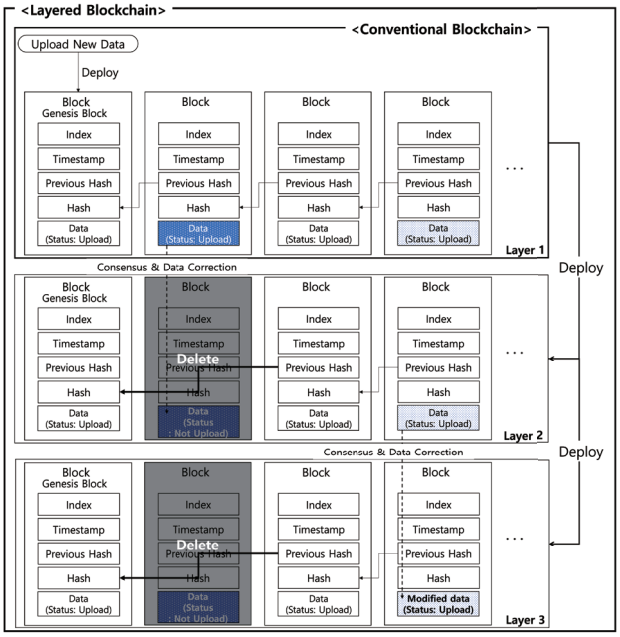


Figure 3. Block diagram of conventional and layered blockchain methods

4.2 Evaluation Results

4.2.1 Latency

In technological applications, latency evaluation is crucial because the proposed method should guarantee low latency to the user. Therefore, we evaluated whether the proposed models could guarantee a latency performance equivalent to a conventional blockchain. The latency evaluation simulation was based on subsequent assumptions. Latency is the time required for data to be uploaded to a blockchain. A conventional blockchain measures the time until data are uploaded to the blockchain. Layered and selective layered blockchains measure the time until the data are uploaded to the top layer. We measured the time to process 1,000 transactions during a round of simulations and calculated the average time of the 1,000 rounds. Table 5 details the seven test cases with distinct sensitivity data ratios. Case 1 had the lowest data sensitivity and processed thousands of public-level transactions. Case 7 exhibited the highest data sensitivity and processed thousands of high-level transactions.

Table 5. Test cases with sensitive data ratios

	Public	Low	Moderate	High
Case 1	100%	0%	0%	0%
Case 2	50%	50%	0%	0%
Case 3	0%	100%	0%	0%
Case 4	33%	33%	33%	0%
Case 5	100%	0%	100%	0%
Case 6	25%	25%	25%	25%
Case 7	0%	0%	0%	100%

Figure 4 depicts the performance evaluation results of the layered blockchain, the selectively layered blockchain, and the conventional blockchain in Cases 1 to 7. The x-axis lists Cases 1–7, and the y-axis denotes the latency until the transaction is uploaded to the block. In Case 1, the conventional, layered, and selectively layered blockchains required 0.47, 9.96, and 0.66 ms, respectively. Public-level transactions are non-sensitive data transmitted through a single-layered blockchain. Therefore, the latency of a selective-layered blockchain is similar to that of a conventional blockchain. In Case 7, the conventional, layered, and selective-layered blockchains required 0.49, 9.99, and 9.97 ms, respectively. High-level transactions are the most sensitive types of data in a four-layer blockchain. Therefore, the latency of the selective-layered blockchain is similar to that of the layered blockchain. The latency does not change according to the data sensitivity of the conventional and layered blockchains. However, the latency of the selective-layered blockchain increases as the data sensitivity increases. Generally, the conventional blockchain was the most instantaneous, whereas the layered blockchain was the most lagging. The conventional blockchain is fast; however, data correction is impossible, whereas the layered blockchain is data-correctable but slow. Selective-layered blockchains can be a suitable option in an environment where the sensitivity level of the data changes.

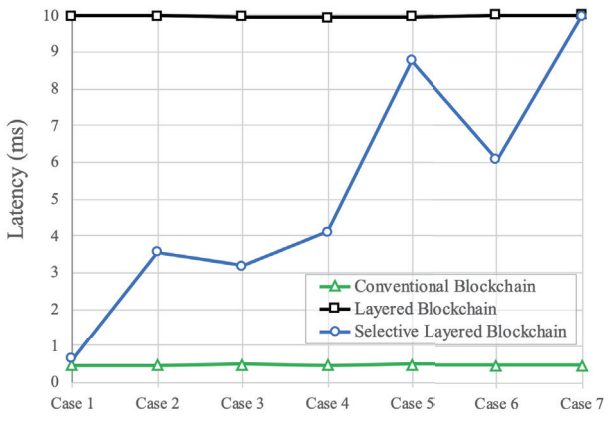


Figure 4. Latency evaluation results of conventional, layered, selective layered blockchain methods

4.2.2 Security

The security evaluation simulation was based on subsequent assumptions. The privacy leakage score represents the degree of data leakage. The degree of data leakage increases as the data sensitivity increases; it is exposed to more users. The privacy leakage score formula for an data transaction is as follows:

$$\begin{aligned} &Privacy\ leakage\ score \\ &= sensitivity\ level\ \times\ scale\ of\ layer. \end{aligned} \tag{2}$$

The sensitivity level is the weight of the transaction sensitivity. The sensitivity level was set as an integer based on the data sensitivity. Public, low, moderate, and high levels were set to 0, 1, 2, and 3 points, respectively. The layer's scale is the weight of the number of nodes exposed to transactions. The layer's scale ranges from 0 to 1, depending on the percentage of total member participation. For example, the scale of the conventional layer equals one, and that of layer 1 of the four-layered blockchain equals 0.125.

Security refers to the degree of privacy leakage in a blockchain. The security can be calculated based on the privacy leakage score. The security formula is as follows:

$$\begin{aligned} &Security = \\ &\sum_{i=1}^{n} (error\ transaction_i\ \times\ privacy\ leakage\ score_i). \end{aligned} \tag{3}$$

Variable n is the number of data transactions with privacy leakage, and i is the number of transactions. Security appends the privacy leakage score of the invalid privacy transactions uploaded to the blockchain. The security of a conventional blockchain can omit the layer's scale because it equals one. For example, if 25 public-level, 30 low-level, 20 moderate-level, and 10 high-level transactions were uploaded to a conventional blockchain, the security score would be 100.

The security-to-latency ratio indicates the effectiveness of preventing data leakage. The security-to-latency ratio is presented as follows:

$$Security\ to\ latency\ ratio = \frac{security}{latency}. \tag{4}$$

This evaluation compared the security of conventional, layered, and selective layered blockchains by measuring the privacy leakage score. Conventional blockchains cannot correct data transactions with privacy leakage. However, layered and selectively layered blockchains can correct data transactions with privacy leakage. We measured the privacy leakage score from when a transaction was issued until the block was uploaded. We measured the time to process 1,000 transactions during a round of simulations and calculated the average time of the 1,000 rounds. During a single round, we issued 1,000 transactions. The data-sensitivity level of the transaction was set randomly.

Figure 5 depicts the security and security-to-latency ratio (SLR) evaluation results for the layered, selective layered, and conventional blockchains. The security values of the conventional, layered, and selectively layered blockchains were 750.7, 516.5, and 334.8, respectively. The corresponding SLR values were 469, 50.1, and 41.3, respectively. Conventional blockchain has a high score because data cannot be corrected. The layered and selective-layered blockchains demonstrated lower scores than conventional blockchains. The security of the selectively layered blockchain is 2.2 times higher than that of the conventional blockchain and 1.5 times higher than that of the layered blockchain. Moreover, the SLR of the selectively layered blockchain is 11.3 times higher than that of the conventional blockchain and 1.2 times higher than that of the layered blockchain.

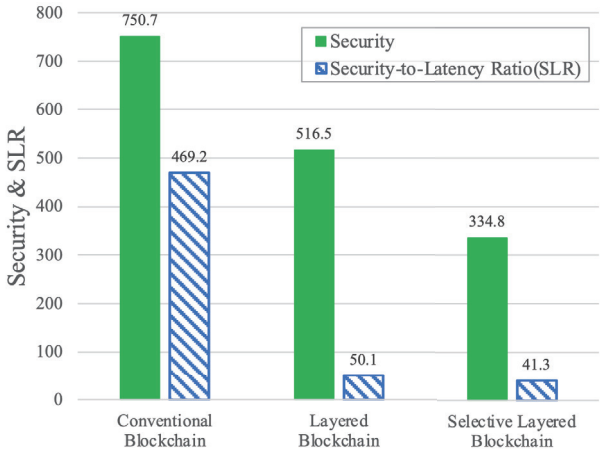


Figure 5. Security evaluation results of conventional, layered, selective layered blockchain methods

4.2.3 Space Complexity

We refer to space complexity as the concept presented in Ahmad's publication [29]. Space complexity indicates the size of the storage used by the blockchain. The storage used by a blockchain can be calculated as *the number of transactions* × *the number of nodes*. In space complexity, *the number of transactions* can be considered because of *the number of transactions* ≫ *the number of nodes*.

Table 6 details the space complexities of the conventional, layered, and selective layered blockchains. The space complexity of the conventional blockchain is $O(t)$. The variable t is the total number of transactions in the blockchain. The space complexity of the conventional blockchain is determined by the size of t. The space complexity of the layered blockchain is $\sum_{i=0}^{l} O(t_i)$. l is the number of layers, and t_i is the number of transactions in layer i. The space complexity of the layered blockchain increases as a function of the number of layers. The number of layers equals the number of data sensitivity classification criteria. Therefore, this number is likely small (comprising a value between 0 and 4). The space complexity of the selectively layered blockchain is $\sum_{i=0}^{l} O(t_i) \times s_i$, and s_i is the proportion of transactions with a data sensitivity level i. Level 0 denores the public level, and level 3 signifies the high level. s_i has a value between 0 and 1. Due to the weight s_i, the space complexity of the selectively layered blockchain is smaller

than that of the layered blockchain. The selectively layered blockchain decreases the number of layers it goes through as it has a lower data sensitivity. Therefore, the space complexity decreases as the data sensitivity decreases.

Table 6. Space complexity of the conventional, layered, selective layered blockchain methods

	Conventional blockchain	Layered blockchain	Selective layered blockchain
Space complexity	$O(t)$	$\sum_{i=0}^{l} O(t_i)$	$\sum_{i=0}^{l} O(t_i) \times s_i$

5 Conclusion

This study investigated a layered blockchain architecture with a data-correction function for managing massive amounts of information without compromising privacy and the right to be forgotten in systems such as mobile IoT networks. Blockchain has an irreversible property, wherein data cannot be reversed once stored in the ledger; this help to defend against data falsification attacks in a P2P environment without a reliable intermediary. The layered blockchain method proposed herein differentially guarantees the right to be forgotten based on a multilayer structure with differential data disclosure ranges by adding a blockchain network as a buffer layer. The right to be forgotten and space complexity are verified by conceptually comparing the proposed blockchain with a conventional blockchain. Performance and security were also evaluated based on simulations to enable comparisons with conventional blockchains. Based on the performance evaluation, the conventional blockchain was the most instantaneous, and the layered blockchain was the most lagging. In the security evaluation, the SLR of the selectively layered blockchain was 11.3 times higher than that of the conventional blockchain and 1.2 times higher than that of the layered blockchain. The space complexity of the layered blockchain increased based on the number of layers compared to the space complexity of a conventional blockchain. In addition, the space complexity of the selectively layered blockchain approaches that of the layered blockchain when the data sensitivity is increased and the space complexity of the conventional blockchain when the data sensitivity is decreased. However, the proposed three-layered blockchain architecture faces several challenges. First, the fields of potential commercial applications were not discussed; thus, research on the utilization of layered blockchain in actual industries should be conducted. Second, the privacy of the layered blockchain was evaluated; however, confidentiality and integrity were not assessed. Therefore, research has to be conducted on internal and external threats and countermeasures that target layered blockchains. Third, a selective-layered blockchain methodology that flexibly manages data according to data sensitivity has been proposed; however, complexity and scalability remain challenging. In future studies, we will investigate a large-scale network environment that transmits and receives considerable data and evaluate the system performance and security dependency on the number of layers. We will also study an extension model with various layered blockchain applications to reduce the structural complexity of the proposed model and expand its scalability.

Acknowledgement

This work was supported by Sungshin Women's University Research Grant H20200081.

Sun-woo Yun and Eun-young Lee contributed equally to this work.

References

[1] S. Nakamoto, Bitcoin: A Peer-to-peer Electronic Cash System, *Decentralized Business Review*, p. 21260, August, 2008.

[2] P. Dutta, T. M. Choi, S. Somani, R. Butala, Blockchain Technology in Supply Chain Operations: Applications, Challenges and Research Opportunities, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 142, Article No. 102067, October, 2020.

[3] M. Javaid, A. Haleem, R. P. Singh, S. Khan, R. Suman, Blockchain Technology Applications for Industry 4.0: A Literature-based Review, *Blockchain: Research and Applications*, Vol. 2, No. 4, Article No. 100027, December, 2021.

[4] I. Al_Barazanchi, A. Murthy, A. A. Al Rababah, G. Khader, H. R. Abdulshaheed, H. T. Rauf, E. Daghighi, Y. Niu, Blockchain-Technology-based Solutions for IoT Security, *Iraqi Journal for Computer Science and Mathematics*, Vol. 3, No. 1, pp. 53-63, January, 2022.

[5] C. Wirth, M. Kolain, Privacy by Blockchain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data, *Proceedings of 1st ERCIM Blockchain Workshop 2018, European Society for Socially Embedded Technologies (EUSSET)*, Amsterdam, Netherlands, 2018, pp. 1-7.

[6] L. Moerel, Blockchain & Data Protection … and Why They are not on a Collision Course, *European review of private law*, Vol. 26, No. 6, pp. 825-851, December, 2018.

[7] N. Sabah, A. Sagheer, O. Dawood, Survey: (Blockchain-based Solution for Covid-19 and Smart Contract Healthcare Certification), *Iraqi Journal for Computer Science and Mathematics*, Vol. 2, No. 1, pp. 1-8, January, 2021.

[8] M. G. Han, D. K. Kang, Toward Multiple Layered Blockchain Structure for Tracking of Private Contents and Right to be Forgotten, In: D. Singh, N. Rajput (Eds.), *Blockchain Technology for Smart Cities*, Springer, 2020, pp. 99-114.

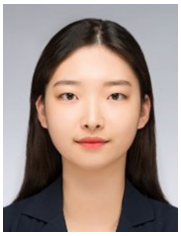
[9] I. C. Lin, T. C. Liao, A Survey of Blockchain Security Issues and Challenges, *International Journal of Network Security*, Vol. 19, No. 5, pp. 653-659, September, 2017.

[10] A. Begum, A. Tareq, M. Sultana, M. Sohel, T. Rahman, A. H. Sarwar, Blockchain Attacks, Analysis and a Model to Solve Double Spending attack, *International*

Journal of Machine Learning and Computing, Vol. 10, No. 2, pp. 352-357, February, 2020.
[11] Q. Feng, D. He, S. Zeadally, M. K. Khan, N. Kumar, A Survey on Privacy Protection in Blockchain System, *Journal of Network and Computer Applications*, Vol. 126, pp. 45-58, January, 2019.
[12] D. Wang, J. Zhao, Y. Wang, A Survey on Privacy Protection of Blockchain: The Technology and application, *IEEE Access*, Vol. 8, pp. 108766-108781, May, 2020.
[13] D. López, B. Farooq, A Multi-Layered Blockchain Framework for Smart Mobility Data-markets, *Transportation Research Part C: Emerging Technologies*, Vol. 111, pp. 588-615, February, 2020.
[14] R. Dennis, G. Owenson, B. Aziz, A Temporal Blockchain: A Formal Analysis, *2016 International Conference on Collaboration Technologies and Systems (CTS)*, Orlando, FL, USA, 2016, pp. 430-437.
[15] G. Zyskind, O. Nathan, A. S. Pentland, Decentralizing Privacy: Using Blockchain to Protect Personal Data, *2015 IEEE Security and Privacy Workshops*, San Jose, CA, USA, 2015, pp. 180-184.
[16] Y. Zhu, Y. Qin, G. Gan, Y. Shuai, W. C. C. Chu, Tbac: Transaction-based Access Control on Blockchain for Resource Sharing with Cryptographically Decentralized Authorization, *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Tokyo, Japan, 2018, pp. 535-544.
[17] I. G. Lee, J. H. Nam, S. J. Lee, Method and Apparatus for Managing Subject Data based on Blockchain, *U.S. Patent*, Washington, USA, No. 11,475,437, May, 2019.
[18] M. Firdaus, K. H. Rhee, On Blockchain-enhanced Secure Data Storage and Sharing in Vehicular Edge Computing Networks, *Applied Sciences*, Vol. 11, No. 1, Article No. 414, January, 2021.
[19] S. Farshid, A. Reitz, P. Roßbach, Design of a Forgetting Blockchain: A Possible Way to Accomplish GDPR Compatibility, *Proceedings of the 52nd Hawaii International Conference on System Sciences*, Hawaii, USA, 2019, pp. 7087-7095.
[20] N. Al-Zaben, M. M. Hassan Onik, J. Yang, N. Y. Lee, C. S. Kim, General Data Protection Regulation Complied Blockchain Architecture for Personally Identifiable Information Management, *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, Southend, United Kingdom, 2018, pp. 77-82.
[21] A. Bayle, M. Koscina, D. Manset, O. Perez-Kempner, When Blockchain Meets the Right to be Forgotten: Technology Versus Law in the Healthcare Industry, *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Santiago, Chile, 2018, pp. 788-792.
[22] I. Puddu, A. Dmitrienko, S. Capkun, µchain: How to Forget without Hard Forks, *Cryptology ePrint Archive*, Report No. 2017/106, February, 2017.
[23] E. Politou, F. Casino, E. Alepis, C. Patsakis, Blockchain Mutability: Challenges and Proposed Solutions, *IEEE Transactions on Emerging Topics in Computing*, Vol. 9, No. 4, pp. 1972-1986, October-December, 2021.
[24] P. Kumar, G. P. Gupta, R. Tripathi, Tp2sf: A Trustworthy Privacy-preserving Secured Framework for Sustainable Smart Cities by Leveraging Blockchain and machine learning, *Systems Architecture*, Vol. 115, Article No. 101954, May, 2021.
[25] P. Kumar, R. Kumar, G. P. Gupta, R. Tripathi, G. Srivastava, P2tif: A Blockchain and Deep Learning Framework for Privacy-preserved Threat Intelligence in Industrial IoT, *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 9, pp. 6358-6367, September, 2022.
[26] T. Li, H. Wang, D. He, J. Yu, Blockchain-based Privacy-preserving and Rewarding Private Data Sharing for IoT, *IEEE Internet of Things Journal*, Vol. 9, No. 16, pp. 15138-15149, August, 2022.
[27] ISO/IEC JTC 1/SC 27, *Iso/Iec 27001:2022 Information Security Management Systems*, Edition 3.0, October, 2022.
[28] National Institute of Standards and Technology, *Standards for Security Categorization of Federal Information and Information Systems*, NIST FIPS Report No. 199, February, 2004.
[29] A. Ahmad, M. Saad, L. Njilla, C. Kamhoua, M. Bassiouni, A. Mohaisen, Blocktrail: A Scalable Multichain Solution for Blockchain-based Audit Trails, *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019, pp. 1-6.

Biographies

Sun-Woo Yun received her B.S. degree in convergence security engineering from Sungshin Women's University, Seoul, Korea, in 2020, and her M.S. degree in future convergence technology engineering from Sungshin Women's University, Seoul, Korea, in 2022. She has authored/coauthored 6 papers in the area of convergence/information security.



Eun-Young Lee received her B.S. degree in convergence security engineering from Sungshin Women's University, Seoul, Korea, in 2021, and her M.S. degree in future convergence technology engineering from Sungshin Women's University, Seoul, Korea, in 2023.

Il-Gu Lee received PhD degree in the Graduate School of Information Security in Computer Science & Engineering Department from KAIST in 2016. He is a professor at the Department of Convergence Security Engineering, Sungshin Women's University, Seoul, Korea. Before joining SWU in March 2017, he was with the ETRI as a senior researcher from 2005 to 2017.