Wireless Communication and Digital Identity Technologies for Secure Financial Services and User Authentication

Sohyun Park

Department of Future Convergence
Technology Engineering
The Graduate School of
Sungshin Women's University

Wireless Communication and Digital Identity Technologies for Secure Financial Services and User Authentication

A Dissertation
Submitted to the
Graduate School of Sungshin University

in partial fulfillment of the requirements
for the degree of
Doctor of Engineering

Sohyun Park
April, 2025

This is to certify that we have examined the
Doctoral Dissertation of
Sohyun Park
Submitted to Department of Future Convergence
Technology Engineering

Approved as to style and content:

Thesis Advisor Il-Gu Lee (signature)

Committee Chairman Seongmin Kim (signature)

Committee Member Il-Gu Lee (signature)

Committee Member Yeon-sup Lim (signature)

Committee Member Kyongjin Kim (signature)

Committee Member Jung Hoon Lee (signature)

The Graduate School of
Sungshin Women's University

ABSTRACT

Wireless Communication and Digital Identity Technologies for Secure Financial Services and User Authentication

Sohyun Park
Department of Future Convergence Technology Engineering
Graduate School of
Sungshin University

With the rapid advancement of information and communication technologies, digital transformation is actively reshaping various sectors, including finance, healthcare, and government services. In the financial sector, traditional payment methods based on physical cards have been progressively replaced by contactless systems utilizing short-range communication technologies such as near-field communication (NFC). Additionally, the development of self-sovereign identity (SSI) technologies has facilitated the adoption of mobile identification systems as alternatives to physical IDs, whereas data-driven financial services such as MyData have enabled user-centric control of personal information.

However, the digitalization of financial services has introduced growing security threats, particularly in wireless communication and user authentication processes. This dissertation proposes a Wi-Fi-based financial payment system that employs physical layer key generation (PLKG) to overcome the distance limitations of conventional financial technologies and to address vulnerabilities in the communication process. Furthermore, a

digital identity protection mechanism is presented to mitigate the risk of personal information leakage during the identification and authentication stages. By integrating security measures at both the physical and logical layers, the proposed system enhances the reliability and security of communication and user management in digital financial services.

Contents

List of Tables

List of Figures

Ⅰ. Introduction

Driven by advancements in information and communication technologies, digital transformation has actively reshaped various sectors, including finance, healthcare, and government services. In the financial domain, conventional payment systems based on physical cards have been progressively replaced by contactless IC card-based technologies that utilize short-range wireless communication methods, such as near-field communication (NFC). This technology enables users to make offline payments using a mobile device preloaded with registered card information instead of carrying a physical card. NFC is a short-range wireless communication technology widely used not only in financial payment systems but also in access control, authentication, and identification systems. Although its limited communication range—only a few centimeters—offers inherent resistance to eavesdropping and related attacks, it significantly compromises user convenience.

The increased digitization of industries such as finance has highlighted the importance of securing the data transmitted and managed wirelessly and online. In particular, the leakage of data and personal information during user and device identification and authentication has become a critical concern.

In an NFC-based payment system, users register their card information and a pre-issued payment token to generate a unique digital identity (ID), enabling secure transactions without exposing sensitive information. Similarly, in financial services such as MyData, application programming

interface (API) tokens are issued based on the user's digital ID to facilitate identification and authentication. However, because these digital IDs are fixed and immutable, their exposure can lead to privacy breaches and enable malicious user tracking, thereby posing a significant security threat.

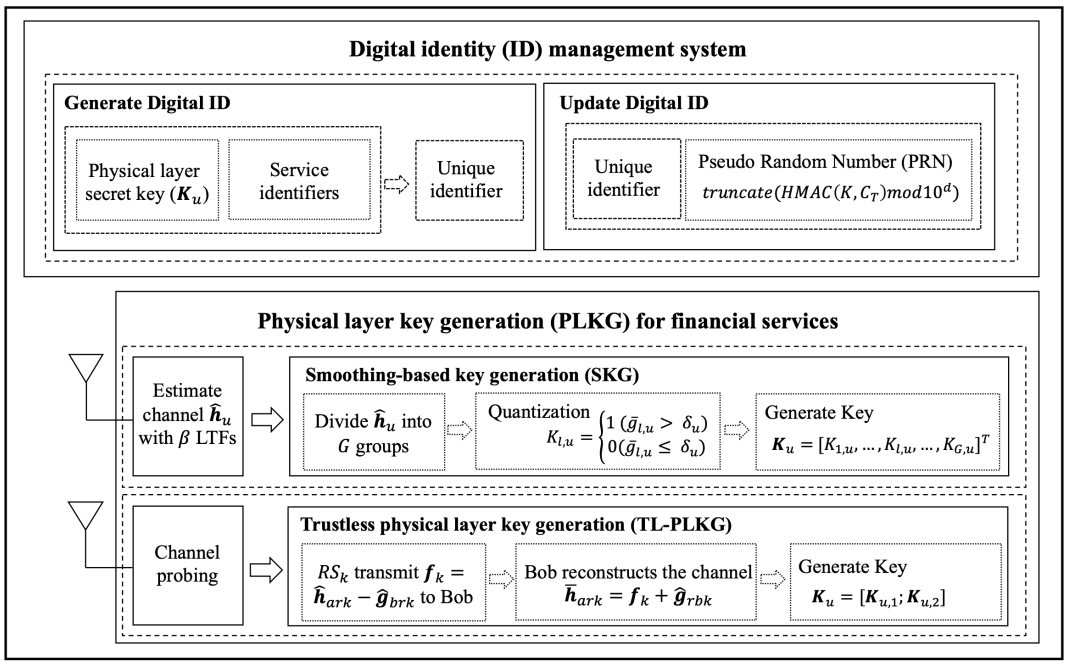

FIGURE 1.1 Overall system architecture of the proposed secure financial transaction framework.

In this dissertation, we propose a wireless communication and digital ID management system that aims to enhance the security of financial payment and user authentication services. Conventional communication technologies used for financial transactions are based on short-range wireless methods, which limit user convenience. Moreover, traditional digital ID systems for identifying users and devices rely on fixed identifiers, making them vulnerable to secondary damage caused by

information leakage and limiting their compatibility across multiple services.

To adrress these limitations, this dissertation proposes a secure financial transaction system that leverages long-range communication technologies, such as Wi-Fi, to minimize the risk of identity leakage by enabling the dynamic generation of digital IDs. As illustrated in Fig. 1.1, the system architecture integrates a smoothing-based key generation (SKG) scheme and a trustless physical layer key generation (TL-PLKG) algorithm for physical layer security, along with a digital ID management system for logical security. When generating a digital ID for user or device identification, a physical layer secret key $\boldsymbol{K}_u$ (where $u \in \{\text{Alice}, \text{Bob}\}$), derived from SKG and TL-PLKG, is processed into a unique identifier. This digital ID is periodically updated based on a pseudo-random number (PRN), thereby preventing secondary damage caused by potential ID exposure.

Chapter 2 introduces a Wi-Fi-based wireless payment system designed to overcome the range limitations of traditional NFC-based systems. Wi-Fi, as a widely adopted wireless communication standard operating in unlicensed bands (2.4, 5, and 6 GHz), offers higher data rates and broader coverage than NFC. By integrating physical layer security techniques based on physical layer key generation (PLKG), this study enhances the security of Wi-Fi communication in financial transactions. The SKG algorithm for PLKG introduced in this chapter enables two legitimate users, Alice and Bob, to repeatedly transmit β long training fields (LTFs) to enhance the received signal-to-noise ratio (SNR) of the LTFs. This

improves the key mismatch rate by grouping and smoothing the subcarriers.

Chapter 3 extends SKG to relay-assisted wireless networks, further improving the coverage and scalability of wireless financial transaction technologies. The TL-PLKG scheme introduced in this chapter is a key generation method designed for untrusted relay network systems. It enables two legitimate users to securely share a key without exposing channel or key information to the relay nodes. The relay-based PLKG scheme enables secure communication while further extending the communication coverage of wireless technologies for financial transactions.

Chapter 4 presents a digital identity protection mechanism at the software and logical layers to address privacy risks associated with static digital IDs. Specifically, a scheme for dynamically updating and synchronizating digital IDs is proposed to mitigate the risk of personal information leakage in financial service applications.

Ⅱ. A Secure Wireless Payment System with Physical Layer Key Generation

1. Introduction

Wireless communication technology is essential across all industries, including information and communication technology, networking, healthcare, and smart homes (Alsabah et al., 2021; Khanh et al., 2022). Notably, the financial sector leverages near-field communication (NFC)-based wireless technology to facilitate contactless payments, often without requiring physical cards (Kulkarni, 2021; Ahamad, 2021). Additionally, NFC is used in applications such as transportation systems, access control, and smart device management. However, unlike other wireless communication technologies such as Wi-Fi and Bluetooth, which can transmit signals over tens of meters, NFC operates within an extremely short range of approximately 5 cm. Its key advantages include ultra-low power consumption and fast, efficient communication attributed to minimal scanning and pairing times.

In applications demanding robust security, such as financial payments and access control, NFC typically incorporates a separate secure element or security chip to handle the encryption and tokenization of sensitive information. Despite this, NFC itself does not inherently support payload encryption, primarily relying on its short operational range as a basic security measure. This reliance makes it susceptible to various attacks, including eavesdropping, spoofing, and replay attacks (Alrawad et al., 2023; Dai et al., 2024; Lee et al., 2021). Furthermore, its strictly limited

communication range can impede usability and convenience. Therefore, there is a clear need for a low-power, high-efficiency, and high-throughput wireless communication technology that can enhance both the security and operational range of NFC.

Wi-Fi stands as the most prevalent and representative standard for wireless communication. Adhering to the latest IEEE 802.11be standard, Wi-Fi operates across the 2.4, 5, and 6 GHz bands and supports channel bandwidths up to 320 MHz. It further accommodates features such as 4096-QAM modulation and up to 16 spatial streams, maximizing both spectral efficiency and data throughput (Abdalhafid et al., 2024; Galati-Giordano et al., 2024; Garcia-Rodriguez et al., 2021). Furthermore, Wi-Fi incorporates strong security mechanisms, notably offering default payload encryption via Wi-Fi Protected Access 3 (WPA3). Given these attributes, Wi-Fi represents one of the most advanced wireless technologies in terms of performance, efficiency, and applicability, making it a suitable candidate for overcoming the inherent security and range limitations of NFC.

However, conventional Wi-Fi security protocols, including WPA3, depend on cryptographic algorithms that can be computationally intensive, potentially rendering them unsuitable for power-constrained devices or applications. To mitigate this challenge, recent research has increasingly focused on physical-layer secret key generation (PLKG) techniques, which harness the intrinsic characteristics of wireless channels (Gao et al., 2024; Shaniri et al., 2023; Zhang et al., 2024). PLKG leverages the principles of channel randomness and reciprocity to generate temporary symmetric keys

between communicating parties. Specifically, in time-division duplex (TDD) systems, the wireless channel between two devices exhibits reciprocity within its coherence time and randomness after the coherence time has passed (Xu et al., 2023).

Given that PLKG exploits the channel reciprocity, even a minor distance deviation can result in an entirely independent channel, making it difficult for an eavesdropper to generate the same key. In contrast to NFC's reliance on proximity, Wi-Fi-based PLKG can offer enhanced security by leveraging these unique wireless channel properties over substantially longer distances. In highly secure environments such as wireless payment or access control systems, PLKG can be combined with upper-layer encryption protocols to provide dual-layer protection for the payload. Alternatively, in resource-constrained environments, encryption can be performed using only the physical-layer secret key, thereby minimizing computational overhead and latency.

The security and reliability of PLKG hinge critically on accurate channel estimation between the communicating devices. Typically, physical-layer keys are generated by exchanging symbols between two devices, Alice and Bob, quantizing channel state information (CSI) or received signal strength (RSS) values into binary keys (Usman et al., 2022; Guo et al., 2021). Consequently, higher channel-estimation accuracy between Alice and Bob directly translates to a higher probability of the two legitimate devices generating matching keys. The key mismatch rate (KMR), defined as the proportion of mismatched keys generated by the two parties, serves as a crucial performance metric. Practical PLKG

implementations necessitate a minimal KMR, which requires effective channel-estimation techniques and sufficient channel signal-to-noise ratio (SNR).

Within the IEEE 802.11 standard, LTF symbols embedded in the packet preamble are specifically designed for channel estimation (Kazaz et al., 2021). These symbols are often transmitted repetitively by allowing the receiver to combine multiple observations, thereby effectively enhancing the received SNR per subcarrier. This improved estimation contributes to significant channel accuracy and, consequently, a reduced KMR in PLKG.

Prior research on improving PLKG has explored various techniques, including applying error correction codes to rectify mismatched key bits, developing novel key quantization schemes, and incorporating auxiliary elements such as reconfigurable intelligent surfaces to enhance KMR. However, these methods tend to be complex, requiring additional computational resources, communication overhead, or specialized hardware (Lu et al., 2021).

Furqan et al. (2020) proposed the Indices-based key generation (IKG) algorithm, which groups the subcarriers of an orthogonal frequency division multiplexing (OFDM) symbol into subgroups and selects only the top n subcarriers exhibiting the highest channel impulse response magnitudes for key generation. The indices of these selected subcarriers are then quantized into binary bits using a predefined lookup table. Although IKG can marginally improve the KMR by focusing on high-SNR subcarriers, it inherently discards channel information from the unselected subcarriers, resulting in information loss.

To address this issue, this study proposes a smoothing-based key generation (SKG) scheme, wherein keys are generated using the subcarriers of the LTF symbol. These subcarriers are first divided into m subgroups, and the average channel gain value within each subgroup is then used for key generation. By utilizing the average gain, this approach incorporates information from all subcarriers, thus mitigating the information loss observed in IKG. Furthermore, averaging adjacent subcarriers exploits frequency diversity, which can further reduce the KMR.

In summary, this study enhances the probability of key agreement between two devices by leveraging the repeated LTF symbols from the IEEE 802.11 preamble for improved channel estimation. It further reduces the KMR by introducing the SKG technique for OFDM-based systems. It also evaluates the key leakage rate (KLR), quantifying the probability of key leakage to an eavesdropper as a function of distance. Finally, the overall performance and security trade-off of the proposed scheme are evaluated using secrecy throughput, a metric designed to capture the interplay between data rate (potentially reduced by LTF repetition), KMR, and KLR.

The primary contributions of this study are as follows:

1) A novel physical-layer security scheme based on Wi-Fi PLKG is proposed to overcome the coverage limitations of conventional short-range communication technologies such as NFC. Additionally, the generated physical-layer secret key can be employed to encrypt header information, such as medium access control (MAC) addresses or destination IDs

within the signal field of the physical layer convergence protocol header, thereby strengthening the security of both the communication system and the transmitted data, particularly in FinTech applications.

2) Repeating the signal in both the frequency and time domains enhances the reception performance of the preamble, thereby enhancing the KMR-generation performance of the PLKG scheme. Additionally, a mechanism to further improve the KMR by excluding poor-quality subcarriers from the SKG process is proposed.

3) A new evaluation metric, secrecy throughput, is introduced to comprehensively assess PLKG schemes by simultaneously considering KMR, KLR, and effective data throughput. Additionally, a method for optimizing the performance and security of the proposed scheme is presented.

The remainder of this chapter is organized as follows: Section 2 reviews related work on conventional PLKG techniques. Section 3 details the proposed SKG method and system model. Section 4 presents simulation results and corresponding analysis to evaluate the performance and security of the proposed technique. Section 5 discusses these results and outlines future research directions, and Section 6 concludes the paper.

TABLE 2.1
Summary of Key Techniques, Contributions, and Limitations of PLKG Studies

Ref.	Year	Parameters	Contributions	Limitations
Zhang et al. (2019)	2019	RSS-based PLKG	- Proposed an authentication and RSS-based key generation scheme using PLS for IoT environments. - Introduced user authentication via RF fingerprinting. - Presented a security framework integrating key generation and authentication.	- Low entropy of RSS - High sensitivity to environmental changes. - Limited security performance.
Lin et al.	2020	RSS-based PLKG	- Applied wavelet shrinkage to denoise RSS values. - Enhanced BDR and BGR using MCQSG algorithm.	- Low entropy of RSS - High sensitivity to environmental changes. - Limited security performance.
Furqan et al.	2020	CFR-based PLKG	- Proposed IKG algorithm using subcarrier positions in OFDM. - Achieved up to 100% improvement	- Discards subcarriers with low channel gain. - Suffers from information loss.

		in key generation rate compared to traditional methods.	
Assaf et al.	2023	- Combined CFR-based key generation with physical unclonable functions (PUFs). - Applied artificial fading (AF) to induce channel randomness in low-dynamic environments. - Enabled high-rate key generation even in static channels.	- Requires additional hardware for PUF implementation. - Limited by hardware dependency.
Zhang et al. (2021)	2021	- Improved PLKG performance through optimized power allocation. - Demonstrated up to 72% increase in SKR via power increase.	- Power increase may introduce inter-channel interference. - Potential degradation in overall communication performance.

2. Related Work

PLKG techniques have been actively studied to enhance the security of wireless communication channels by exploiting their inherent physical properties. Early PLKG approaches often employed relatively simple channel metrics, such as RSS.

However, RSS-based methods suffer from limitations, including high sensitivity to environmental fluctuations and a potential for asymmetry between the uplink and downlink channels, which can consequently increase the KMR. As alternatives or complements to RSS, techniques leveraging more detailed channel information—such as channel frequency response (CFR), channel impulse response, phase, and amplitude, collectively known as CSI—have been extensively explored. Furthermore, recent investigations have aimed at improving the secret key rate (SKR) by employing auxiliary technologies such as reconfigurable intelligent surfaces to manipulate channel characteristics and enhance randomness. Additionally, methods focusing on optimizing power allocation to boost key-generation performance have also been actively investigated. Table 2.1 provides a summary of key techniques, contributions, and limitations found in representative PLKG studies.

Zhang et al. (2019) proposed an authentication and key-generation scheme using physical-layer security (PLS) techniques tailored for Internet of Things (IoT) environments. Leveraging the inherent wireless channel properties of randomness and reciprocity, they introduced an RSS-based cryptographic key-generation method for secure communication between IoT devices. They also incorporated user

authentication via radio frequency fingerprinting, presenting a security framework integrating both key generation and authentication. Similarly, Lin et al. (2020) proposed an efficient RSS-based secret key-generation method by employing a wavelet shrinkage technique to denoise the collected RSS values and using the modified channel quantization with a single-guard-band algorithm to improve the bit disagreement and bit generation rates. While both studies proposed efficient PLKG techniques, particularly suitable for low-power environments, they share the fundamental limitation associated with RSS: its relatively low entropy and high sensitivity to environmental changes, which restrict the achievable security level.

In contrast to methods relying solely on signal amplitude or phase, Furqan et al. (2020) presented a notable technique utilizing subcarrier positions within an OFDM system for key generation. Their proposed IKG algorithm partitions the OFDM subcarriers into subgroups and selects only those subcarriers exhibiting the highest channel gain within each subgroup for key generation. This method was reported to improve the key-generation rate by up to 100% compared to some conventional PLKG algorithms. Furqan et al. also proposed a joint key generation (JKG) algorithm, combining keys derived from the IKG method with those generated by quantizing all subcarriers. However, the primary drawback of IKG is the information loss resulting from discarding subcarriers deemed to have lower channel gain. While the JKG algorithm compensates for this by including the key bits generated from all subcarriers, it may inadvertently increase the KMR owing to potential

inconsistencies introduced by the quantization of lower-quality subcarriers.

Assaf et al. (2023) proposed a PLKG method combining CFR-based key generation with physical unclonable functions (PUFs), enabling high-rate key generation even in environments lacking sufficient channel randomness. In scenarios with near-static or low-variation channels, the authors employed artificial fading techniques to deliberately introduce channel fluctuations, thereby improving key-generation performance. However, PUF necessitates specific hardware design and implementation, introducing additional costs and hardware dependency limitations.

Furthermore, optimizing power allocation has been explored to enhance PLKG performance. Zhang et al. (2021) investigated this aspect and demonstrated that strategically increasing transmission power could significantly enhance the SKR, reporting improvements of up to 72%. They concluded that optimized power allocation leads to better channel estimation and higher receive SNR, thereby enhancing key generation performance. However, a potential drawback of increasing transmission power specifically for key generation is the risk of causing inter-channel interference, which can degrade overall communication system performance.

The SKG algorithm proposed in this study aims to address the information loss limitation identified in the IKG algorithm and leverages repeated LTF symbols to resolve the interference issue caused by power allocation.

3. System Model and the Proposed Method

This outlines a wireless communication system architecture designed to enhance security for applications such as financial transactions by incorporating PLKG. By leveraging the preamble structure within the physical-layer transmission, two legitimate devices can generate shared secret keys. These physical-layer keys can then be used to encrypt signaling information, thereby offering a supplementary layer of security.

Conventional MAC layer encryption typically protects only the frame payload—that is, the actual user data—whereas the header is transmitted in plaintext. This is often done to minimize processing delays caused by encryption and decryption processes and maximize throughput. However, transmitting header fields such as MAC addresses or destination IDs unencrypted exposes the communication to potential attackers who might exploit this information for replay attacks, spoofing, and other malicious activities.

In contrast, PLKG, based on the reciprocity of the wireless channel between legitimate users, facilitates a lightweight symmetric encryption method suitable for encrypting frame headers. Moreover, when used in conjunction with existing upper-layer encryption schemes, PLKG provides an additional security layer, thereby enhancing the overall communication security.

1) Channel estimation with LTFs

As shown in Fig. 2.1(a), Alice and Bob each have a single antenna. They exchange an IEEE 802.11 packet with multiple LTFs and data OFDM symbols over the same frequency channel using a TDD for coherence time, τ. As shown in Fig. 2.1(b), the packet consists of s symbols comprising β LTFs and $s-\beta$ data symbols, where β refers to the number of LTFs. The channel between Alice and Bob is assumed to be a Rayleigh fading channel, represented as $\boldsymbol{h}_{ab} \in \mathbb{C}^{(N\times 1)}$ for the channel from Alice to Bob and $\boldsymbol{h}_{ba} \in \mathbb{C}^{(N\times 1)}$ for the channel from Bob to Alice. N is the number of subcarriers within a symbol.

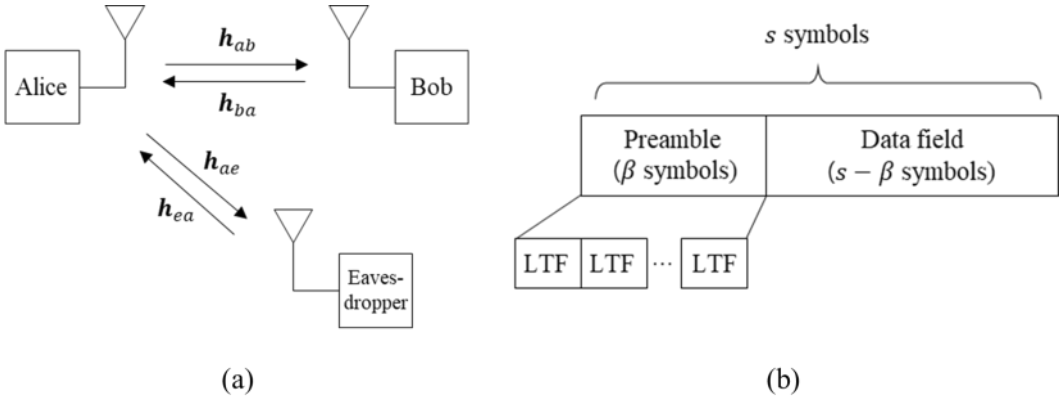

FIGURE 2.1 System model: (a) network configuration and (b) frame structure.

Consider $\boldsymbol{h}_{ae} \in \mathbb{C}^{(N\times 1)}$ as the channel between Alice and the eavesdropper and ρ to be the correlation coefficient between $\boldsymbol{h}_{ab}$ and $\boldsymbol{h}_{ae}$, then $\boldsymbol{h}_{ae}$ can be expressed as $\boldsymbol{h}_{ae} = \rho \boldsymbol{h}_{ab} + \sqrt{(1-\rho^2)}\,\boldsymbol{w}$, where $\boldsymbol{w} \sim CN(\boldsymbol{0}, \boldsymbol{I})$.

When Alice sends a packet to Bob, Bob estimates the channel using the

LTF located in the preamble of the OFDM packet. Similarly, when Bob sends a packet to Alice, Alice estimates the channel using the LTF. Because both exchange packets within coherence time τ, it is assumed that the channel has the reciprocity property (i.e., $\boldsymbol{h}_{ab} = \boldsymbol{h}_{ba}$). The reciprocal channel between Alice and Bob is referred to as h_u.

The LTF received by Alice or Bob is expressed by

$$y_{(c,u)}[k] = h_{c,u}x_{c,u} + z_{c,u}[k], \tag{2.1}$$

where $y_{(c,u)}[k]$ denotes the received value at the c-th subcarrier indices of the k-th LTF symbol received by $u(u \in \{\text{Alice,Bob}\})$ and $c \in \{1,\cdots,\text{N}\}$, $h_{c,u}$ represents the reciprocal channel of the c-th subcarrier between Alice and Bob, and $x_{c,u}[k]$ indicates the transmitted LTF signal. The LTF consists of repeated identical symbols, the sequence of LTF symbols is omitted and simply represented as $x_{c,u}$ in Eq. (2.1). $z_{c,u}[k] \sim CN(0,\sigma^2)$ denotes the additive white Gaussian noise.

2) Key Generation Algorithms

Alice and Bob quantize the channel values of each subcarrier and estimated them using the LTF to generate key bits. The average channel values of 52 subcarriers within a symbol was used as a reference, and subcarrier channel values greater than the average were quantized as 1. By contrast, those less than or equal to the average were quantized as 0, generating the key. Next, 52 key bits were generated using 1-bit quantization.

KMR is the probability that Alice's and Bob's keys do not match during a given sampling period. The KMR of Alice and Bob varies depending on the channel-estimation performance. Alice and Bob can generate a key only if the generated key bits are identical. For example, if KMR is 0.5 out of 10 attempts to generate a key, the key cannot be generated five times.

3) Proposed algorithm

① LTF repetition and SKG

As shown in Fig. 2.2, the proposed SKG algorithm consists of channel estimation with LTFs, subcarrier grouping and smoothing, quantization, and key generation phases.

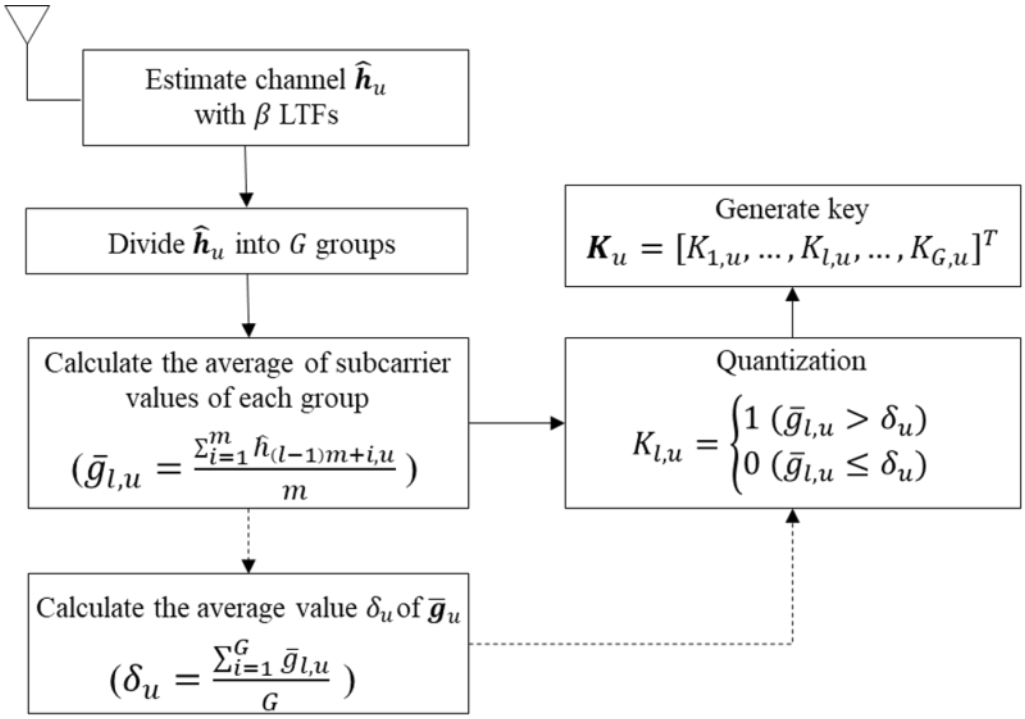

FIGURE 2.2 Key generation algorithm.

1) LTF repetition: In this study, we proposed a method to improve the channel-estimation performance and reduce the KMR by repeating the LTFs in the time domain. The channel value estimated by Alice or Bob with β LTFs using the minimum mean square error channel-estimation technique are expressed as

$$\hat{h}_{c,u} = \frac{x_{c,u}^{*}}{|x_{c,u}|^2 + \sigma^2} \bar{y}_{c,u}, \tag{2.2}$$

$$\bar{y}_{c,u} = \frac{1}{\beta} \sum_{k=1}^{\beta} y_{c,u}[k], \tag{2.3}$$

The SNR increases proportionally with the number of LTF repetitions. By averaging the β LTF channel values, β-times SNR gain can be achieved.

2) Subcarrier grouping/smoothing: The key length decreases by grouping subcarriers and generating keys based on the average channel value within each group; however, the probability of key bit matching between Alice and Bob increases. Assuming that $\hat{\boldsymbol{h}}_u = [\hat{h}_{1,u}, \cdots, \hat{h}_{N,u}] \in \mathbb{C}^{(N \times 1)}$ is the channel between Alice and Bob, the division of $\hat{\boldsymbol{h}}_u$ into G groups, each containing m subcarriers can be expressed as $\hat{\boldsymbol{h}}_u = [\boldsymbol{g}_{1,u}, \cdots, \boldsymbol{g}_{l,u}, \cdots, \boldsymbol{g}_{G,u}]^T \in \mathbb{C}^{(Gm \times 1)}$ with $\boldsymbol{g}_{l,u} = [\hat{h}_{(l-1)m+1,u}, \cdots, \hat{h}_{(lm,u)}] \in \mathbb{C}^{(1 \times m)}$. Here, $\boldsymbol{g}_{l,u}$ refers to the channel vector of the l-th group in $\hat{\boldsymbol{h}}_u$, which consists of m subcarriers. In this study, we defined only cases in which N was divisible by m.

The average value of all elements of $\boldsymbol{g}_{l,u}, \bar{g}_{l,u} = \frac{\sum_{i=1}^{m} \hat{h}_{(l-1)m+i,u}}{m}$ was quantized and used as the key, whose length decreased to $G = \frac{N}{m}$ bits as m increased. However, smoothing the m channel values can reduce the

KMR, and averaging the subcarrier values in $\boldsymbol{g}_{l,u}$ reduces the impact of noise and increases the likelihood that Alice's and Bob's keys matches. Finally, Alice or Bob generates a key using the channel vector $\bar{\boldsymbol{g}}_u = [\bar{g}_{1,u}, \cdots, \bar{g}_{l,u}, \cdots, \bar{g}_{G,u}]^T$.

3) Quantization and key generation: Alice and Bob generated binary key bits $\boldsymbol{K}_u = [K_{1,u}, \cdots, K_{l,u}, \cdots, K_{G,u}]^T$ by quantizing the channel values of subcarriers. If the channel value of a subcarrier was greater than the overall average channel value $\delta_u = \frac{\sum_{i=1}^{G} \bar{g}_{i,u}}{G}$, it was quantized to 1; else, it was quantized to 0.

② Smoothing and filtering-based key generation

To further enhance SKG performance, we proposed the smoothing and filtering-based key generation (SFKG) algorithm, which considers the received signal quality of subcarriers. Specifically, it excludes subcarrier groups with SNR values below a threshold from the key generation process of the SKG algorithm. This technique improves the KMR while minimizing information loss caused by filtering out low-quality subcarriers.

1) Grouping and averaging: Consider $\boldsymbol{\gamma_u}=[\gamma_{1,u},\cdots,\gamma_{N,u}]\in\mathbb{R}^{(N\times 1)}$ as the received SNR of a subcarrier. The average received SNR of the G grouped subcarriers is then expressed as $\bar{\boldsymbol{\gamma}}_u=[\bar{\gamma}_{1,u},\cdots,\bar{\gamma}_{l,u},\cdots,\bar{\gamma}_{G,u}]\in\mathbb{R}^{(G\times 1)}$.

2) Filtering: Among the elements in $\bar{\boldsymbol{\gamma}}_u$, only those with values equal to or greater than the threshold γ_{th} are retained, and subcarriers with poor quality are excluded from the key generation process. The key generated by SFKG algorithm is defined as $\boldsymbol{K}_u=\{\mathrm{K}_{\mathrm{l,u}}|\bar{\gamma}_{\mathrm{l,u}}\geq\gamma_{\mathrm{th}}\}$.

The threshold γ_{th} is a pre-shared value between Alice and Bob and represented as a lookup table. Specifically, it represents the minimum subcarrier SNR required to guarantee the lowest KMR based on the channel's SNR. Alice and Bob share the indices of the subcarrier groups to be excluded, enabling them to filter the same subcarrier groups. Although this index-sharing approach allows an eavesdropper to infer the key length and indices of the filtered subcarrier groups, it does not directly leak the key bit values. The security implications of this index-sharing method, in terms of the KLR, are evaluated in Section 4.

4. Results and Discussion

1) Simulation results

The lengths of the OFDM symbol and cyclic prefix, as well as other parameters, followed the IEEE 802.11 standard. The channel between Alice and Bob was modeled as a multipath channel with an exponentially decaying channel response based on a fixed sampling period of 50 ns. The amplitude of the channel response followed a Rayleigh distribution. The estimated channels between Alice and Bob, $\hat{\boldsymbol{h}}_{ab}$ and $\hat{\boldsymbol{h}}_{ba}$, were modeled as $\boldsymbol{h}_{ab} = \hat{\boldsymbol{h}}_{ab} + \boldsymbol{z}_{ab}$ and $\boldsymbol{h}_{ba} = \hat{\boldsymbol{h}}_{ba} + \boldsymbol{z}_{ba}$, respectively, incorporating channel-estimation errors $\boldsymbol{z}_{ab}$ and $\boldsymbol{z}_{ba}$.

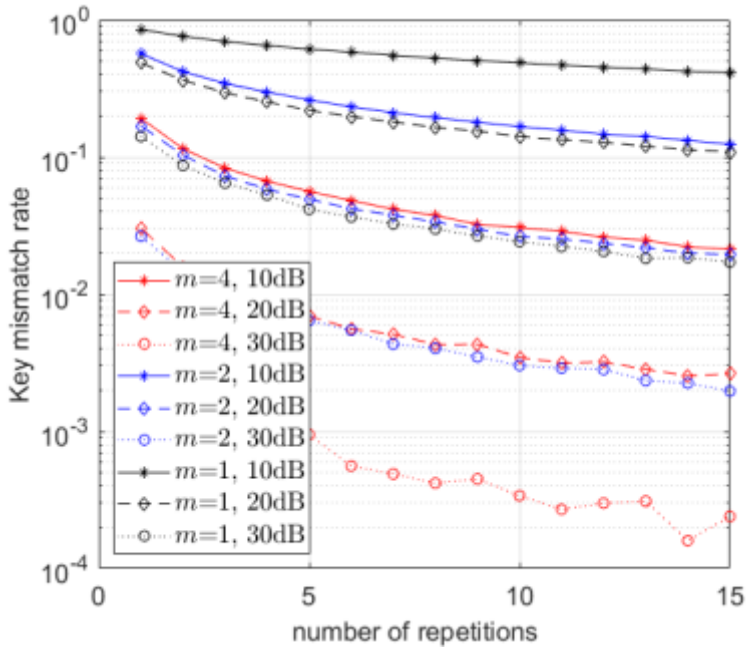

FIGURE 2.3 Key mismatch rate versus the number of LTF repetitions with varying m and the SNR.

Fig. 2.3 shows the KMR according to the number of LTF repetitions (β) when the SNR were 10, 20, and 30 dB and the values of m were 1, 2, and 4, respectively. As the number of LTF repetitions increased in the time domain, the channel-estimation accuracy between Alice and Bob improved, thereby increasing the probability of matching keys. As the number of LTF repetitions increased, KMR decreased. In addition, as m increased, the key length decreased. However, because the key was generated using the average value of m subcarriers, the probability that Alice's and Bob's keys would match increased, resulting in a lower KMR.

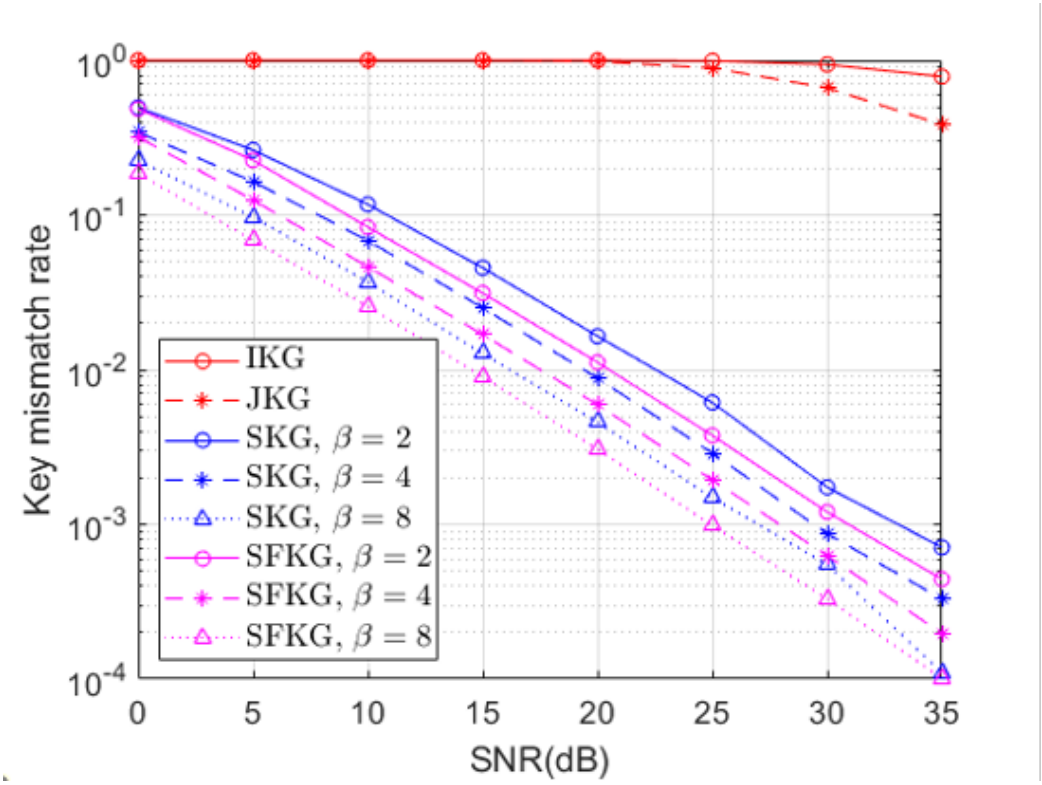

FIGURE 2.4 Key mismatch rate versus SNR (dB) for IKG, JKG, and the proposed SKG and SFKG algorithms.

Fig. 2.4 compares the KMR of IKG, JKG, and the proposed SKG and SFKG algorithms based on varying SNR levels. The KMR performance of the conventional and proposed algorithms was compared under the same conditions as those shown in Fig. 2.4: the LTF symbol, composed of 52 subcarriers, was divided into 13 subgroups (e.g., $m=4$), and in the IKG and JKG algorithms, the key was generated using only the top two values out of the four subcarrier channel values. The evaluation results indicated that the IKG and JKG algorithms did not fundamentally improve channel-estimation performance, resulting in a higher KMR. These algorithms were configured to select the indices of the top two subcarriers within each group of four for key generation. The results highlight that IKG and JKG, which do not inherently improve channel estimation and discard information from lower-gain subcarriers, exhibit significantly higher KMR compared to the proposed methods that utilize all subcarrier information via smoothing. This underutilization of channel resources contributes to their poorer performance. Specifically, at an SNR of 10 dB, the proposed SKG algorithm ($\beta=2$) achieved a KMR reduction of approximately 88.2% compared to that of the IKG algorithm.

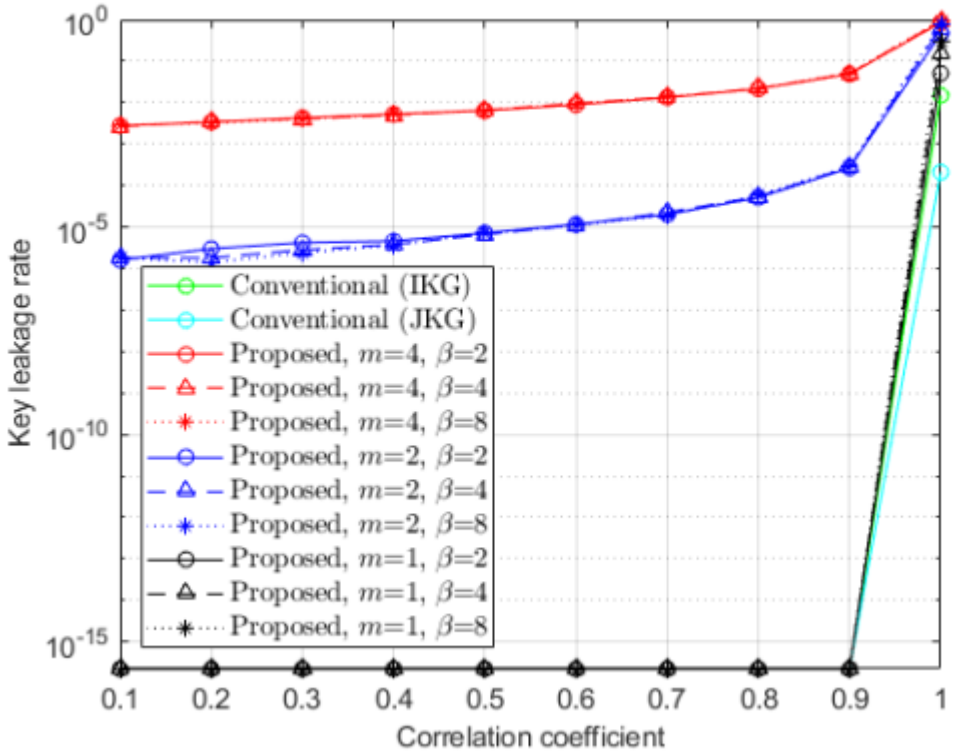

FIGURE 2.5 Key leakage rate versus the number of LTF repetitions varying the SNR levels.

Fig. 2.5 presents the KLR as a function of the spatial correlation coefficient (ρ) between the $\boldsymbol{h}_{ab}$ and $\boldsymbol{h}_{ae}$ channels increased from 0.1 to 1, in 0.1 increments, with the SNR of both $\boldsymbol{h}_{ab}$ and $\boldsymbol{h}_{ae}$ channels set to 10 dB. The IKG and JKG algorithms display a KLR similar to that of the proposed method when $m=1$. However, for the proposed method, the KLR increases as ρ and m increased. Additionally, the number of LTF repetitions and KLR are independent of each other.

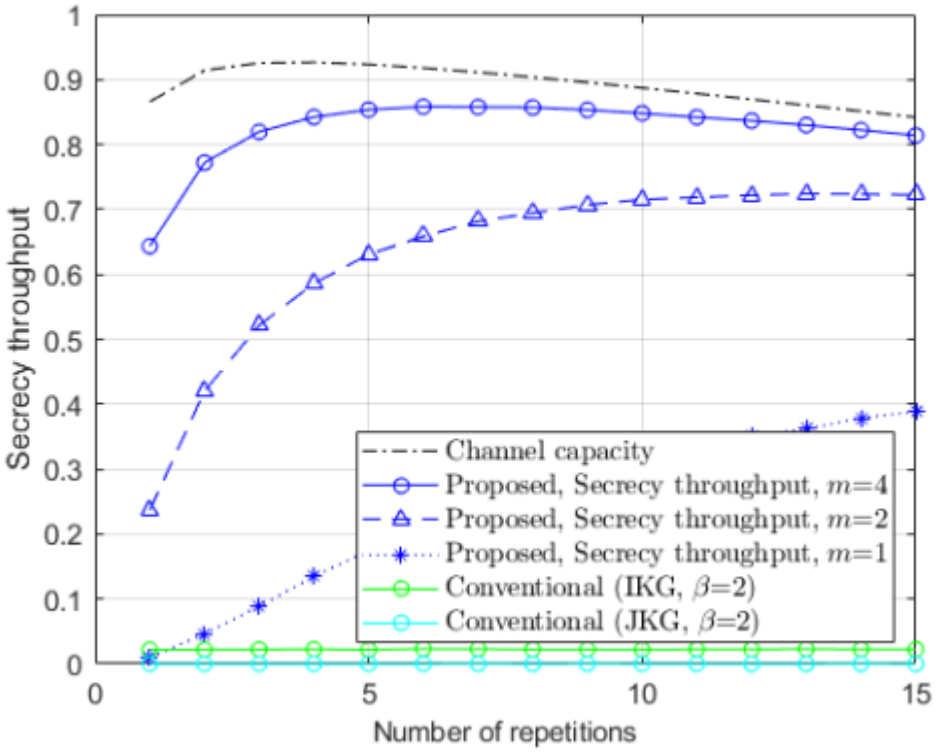

FIGURE 2.6 The secrecy capacity versus the number of LTF repetitions in the frequency domain with varying m.

Fig. 2.6 shows the upper bound of the channel capacity based on the number of LTF repetitions (β) for an SNR of 10 dB and $\rho=0.5$. When the effective SNR after averaging the repeated LTFs is given by $\Gamma=\beta\times SNR$ on a linear scale, the channel capacity (C) and secrecy throughput (C_{sec}) are defined as (Yu et al., 2019; Yu et al., 2020)

$$C = \frac{s-\beta}{s} \times \log_2\left(1 + \frac{\frac{\Gamma}{1+\Gamma}}{1+\frac{1}{1+\Gamma}}\right), \tag{2.4}$$

$$C_{sec} = (1-KLR) \times (1-KMR) \times C. \tag{2.5}$$

The channel capacity can be determined using the channel SNR estimated from the preamble comprising β LTFs. Here, C represents the channel capacity when transmitting the $s-\beta$ data symbols, excluding the preamble, as shown in Fig. 2.1(b). Specifically, C decreases as the number of LTFs increases. C_{sec} represents the effective channel capacity considering the impact of both the KMR and the KLR on the reliability and confidentiality of key generation. KMR decreases as the number of LTF repetitions increases, improving the key-generation performance. However, the amount of data transmitted within a given period decreases as the number of LTF repetitions increases, thereby reducing the channel capacity.

According to the evaluation results, for $s=100$, the KMR of the proposed algorithm decreases as m increases, resulting in the highest C_{sec} when $m=4$ in an environment with a SNR of 10 dB and $m \in \{1,2,4\}$. Additionally, when $m=4$, C_{sec} increases as the number of LTF repetitions increases, reaching its peak at $\beta=6$, before gradually decreased. For an SNR of 10 dB, the proposed method, with $m=4$ and $\beta=6$, showed a 93.7% reduction in KMR compared to the conventional method without additional LTF repetitions or subcarrier grouping, i.e., $m=1$ and $\beta=2$. This demonstrates that the proposed method optimizes C_{sec}, that is,

key-generation performance, security, and data-transmission performance.

Because the IKG and JKG algorithms did not repeat the LTFs more than twice, their C_{sec} values were compared with the LTFs fixed to two. As shown in Fig. 2.5, both IKG and JKG algorithms exhibited extremely low KLRs. However, as shown in Fig. 2.4, this is because no keys are generated; therefore, keys cannot be leaked. Although the KLR of the IKG algorithm converged to 0 at this SNR, it exhibited a high KMR of around 0.99, approximately 20.5 times higher than that of the proposed method with $m=4$ and $\beta=6$. Consequently, compared to the IKG algorithm, the proposed method achieved an approximately 37 times higher C_{sec}, considering both KLR and KMR.

2) Simulation results analysis

A key aspect of this study is the definition of KMR. Some previous studies (Zhang et al., 2019; Furqan et al., 2020) defined KMR as the proportion of mismatched bits among all key bits from the perspective of channel entropy. However, in practical environments, a secret key can only be used if Alice and Bob generate an identical key without any errors. Therefore, this study considered that a key was successfully generated only when all bits were identical, and KMR was defined based on this criterion. Simulation results showed that using this method reduced KMR by 88.2% compared to that with the conventional IKG algorithm, indicating that increasing the LTF repetition improves channel-estimation accuracy, reducing KMR and enabling more precise and efficient key generation.

The proposed SKG/SFKG methods address the limitations of IKG by utilizing information from all subcarriers, rather than discarding those with lower channel gain. By grouping and averaging (smoothing), the algorithms effectively leverage frequency diversity and improve robustness to noise, leading to lower KMR compared to IKG. Furthermore, repeating LTF symbols reduces the amount of data that can be transmitted during the same period, decreasing the overall data throughput. The performance of the proposed method was evaluated using the secrecy throughput metric, which considers both KMR and KLR for key generation performance and throughput for data-transmission efficiency.

The proposed SKG scheme reduces the KMR by 88.2% compared to the conventional IKG algorithm. However, the legacy algorithms, IKG and

JKG, ultimately fail to generate valid keys, resulting in significantly lower KLR than SKG. Nevertheless, when evaluating performance using C_{sec} —a comprehensive metric that considers both KMR and KLR—the optimized SKG algorithm achieves a C_{sec} value that is 37 times higher than that of the conventional schemes.

The optimal LTF repetition and subcarrier grouping parameters, which balance data throughput and key generation performance, were derived. The experiments demonstrated that for the simulated environment (SNR 10 dB), secrecy throughput was optimized when LTF was repeated up to six times and subcarriers were grouped in sets of four. Under these optimized conditions, the proposed method achieved a 37-fold increase in C_{sec} compared to the IKG algorithm.

5. Discussion and Future Work

This study proposed a Wi-Fi-based wireless communication technique incorporating PLS mechanisms as an alternative to NFC technology. Specifically, we introduced the SKG algorithm and its extension, the SFKG algorithm, to enhance the performance of physical-layer secret-key generation.

The SKG algorithm reduces the KMR by grouping and smoothing subcarriers, leveraging channel estimates obtained from the LTF symbols more robustly. However, this grouping process reduces the potential key length. To address this, we assumed a transmission environment using narrower subcarrier spacing, thereby increasing the number of subcarriers within the same bandwidth. This approach preserves key length while improving the KMR.

Future work should focus on implementing and evaluating the proposed SKG and SFKG algorithms in real-world scenarios. More realistic channel models will be considered by incorporating factors such as channel estimation errors and channel variations, to better capture the correlation among Alice, Bob, and Eve and assess the impact on key generation performance. We also plan to model a Wi-Fi-based financial transaction system to evaluate throughput and security performance when SKG and SFKG encryption techniques are applied to both the header and payload of transaction data packets.

Finally, the PLKG-based Wi-Fi payment systems may lead to increased transaction latency compared to traditional payment mechanisms. Future research will evaluate the latency overhead introduced by the proposed

method and develop techniques to reduce it.

6. Conclusion

This study proposed the SKG and SFKG algorithms to enhance the key generation performance of PLKG techniques, aiming to enhance the security of wireless communications. These algorithms demonstrate potential as alternatives to short-range wireless communication technologies, particularly for financial transactions, by addressing its communication-range limitations.. Relying solely on wireless channel reciprocity, both SKG and SFKG offer pathways to fast and lightweight security implementations. Moreover, the inherent spatial sensitivity of wireless channels provides strong security guarantees, ensuring that even slight changes in the environment make it difficult for eavesdroppers to regenerate the same secret key.

Experimental results demonstrated that the proposed SKG algorithm achieves an 88.2% reduction in KMR compared to the conventional IKG algorithm. Furthermore, the SFKG algorithm yielded additional improvement, achieving an additional 31% reduction in KMR relative to SKG. This study also analyzed the trade-off between the channel capacity reduction resulting from LTF repetition and the associated improvements in key generation reliability. Evaluating the SKG algorithm using the secrecy throughput metric, which incorporates KMR, KLR, and channel capacity, revealed that applying an optimal combination of LTF repetition and subcarrier grouping/smoothing resulted in a 37-fold improvement in secrecy throughput compared to the IKG algorithm. As future work, we plan to evaluate the proposed algorithms in real-world environments.

Table of Notations (Chapter II)

Notations	Description
s	The number of data symbols
β	The number of LTF symbols
τ	The coherence time
$\boldsymbol{h}_{ab}$, $\boldsymbol{h}_{ba}$	Channels between Alice and Bob
$\boldsymbol{h}_{ae}$	Channel between Alice and Eve
$\boldsymbol{h}_{u}$	Reciprocal channel between Alice and Bob
u	Legitimate users ($u \in \{\text{Alice}, \text{Bob}\}$)
ρ	Channel correlation coefficient
c	Index of subcarrier
k	Index of LTF symbol
x	Rransmit signal
z	Additive white Gaussian noise
N	The number of subcarriers
G	The number of subcarrier groups
m	The number of subcarriers within a group
l	Index of subcarrier group
$\boldsymbol{K}_{u}$	Quantized secret key of legitimate users
δ	The average channel value
γ	Received SNR
Γ	Effective SNR
C	Channel capacity
C_{sec}	Secrecy throughput

Ⅲ. Physical Layer Key generation in Untrusted Relay Networks

1. Introduction

Wi-Fi is a wireless communication standard that supports device-to-device connectivity in applications such as the Internet of Things (IoT) and smart homes (Szott et al., 2022; Fanari et al., 2024; Abyaneh et al., 2023). Over time, Wi-Fi has evolved to maximize its spectral efficiency and throughput (Yang et al., 2021). In particular, the recently commercialized IEEE 802.11be standard supports communications not only in the 2.4 GHz and 5 GHz bands but also in the higher 6 GHz band, enabling extremely high data rates and large channel capacity with bandwidths up to 320 MHz (Deng et al., 2020; López-Raventós and Bellalta, 2022; Lopez-Perez et al., 2019). However, signals at higher frequencies suffer from low penetration capability, which can lead to the coverage of dead zones in indoor environments, and their propagation range is generally shorter than that of lower frequency bands (Saha, 2021).

In contrast, the IEEE 802.11ah standard, commonly used in smart homes, sensor networks, and industrial IoT, operates in the unlicensed 900 MHz band, allowing communication over a wide coverage range of 1 km or more. Therefore, IEEE 802.11ah has been widely adopted for low-power IoT devices (Ahmed et al., 2022; Tian et al., 2021; Alam et al., 2024). However, its relatively low data rate results in degraded

communication performance.

To address these issues, the upcoming Wi-Fi standard IEEE 802.11bn, which is currently under development, defines reliability as one of its key objectives and includes standardization efforts for relay-based operations to extend coverage (Jeon et al., 2024). Relay technology allows coverage to be extended even in environments where a direct link is not available by using intermediate relay nodes (Yi et al., 2025). Consequently, the relay-assisted Wi-Fi technology can overcome the coverage limitations of conventional Wi-Fi technology and potentially replace IEEE 802.11ah in scenarios where higher data rates are required.

However, relay-based communication poses the potential risk of data leakage through relay nodes (Liu et al., 2023; Sun et al., 2023; Luo et al., 2023). There are two main types of relays: amplify-and-forward (AF) relays, which simply amplify and forward the received signal; and decode-and-forward (DF) relays, which decode the received message and re-encode it before forwarding. During AF relaying, the relay operates only at the signal level, rendering the risk of information leakage relatively low. In contrast, DF relays decode messages, significantly increasing the risk of data exposure through the relay.

DF relays cannot decode encrypted signals. However, they can access the content of unencrypted messages, making them vulnerable from a security standpoint. In environments in which upper-layer encryption protocols are difficult to apply, there is a potential risk of message confidentiality being compromised by the DF relay. Therefore, in untrusted relay environments, new encryption schemes are required to

ensure that neither the secret key nor the message content is exposed to the relay.

Physical layer key generation (PLKG) is a physical layer security technique that generates temporary secret keys by exploiting the channel state information (CSI) of reciprocal wireless channels between two legitimate users. Because PLKG utilizes only physical layer characteristics such as channel frequency response and received signal strength as common randomness for generating symmetric keys, it has gained attention as a lightweight alternative to upper-layer cryptographic techniques (Gao et al., 2024; Xiao et al., 2025; Abdelazeem et al., 2024; Han et al., 2022). This makes it particularly suitable for resource-constrained environments such as IoT environments. However, conventional PLKG schemes assume the existence of a direct link between legitimate users and therefore cannot be directly applied to relay-aided systems. In such cases, a new PLKG scheme is required that enables key generation without revealing any channel- or key-related information to untrusted relays.

In this paper, we propose a trustless physical layer key-generation (TL-PLKG) scheme that enables secret key generation by dividing the frequency band into multiple untrusted relay environments. In addition, we introduce a secret key-sharing algorithm that leverages the reciprocity property of the wireless channel during the key-generation process to securely distribute key information.

Most previous studies on PLKG in relay-assisted environments assumed perfect time synchronization between two legitimate users for key

generation (Aldaghri and Mahdavifar, 2020; Thai et al., 2016; Keshavarzi et al., 2024; Xu et al., 2024). However, achieving such synchronization in real-world scenarios is difficult. In addition, in AF relay schemes, accurately estimating the relay gain for signal separation is prone to errors. By contrast, the proposed TL-PLKG leverages multiple untrusted relays, each of which is responsible for delivering only partial channel information. This design prevents relays from accessing the knowledge of the entire channel. By allocating only subsets of the total bandwidth and subcarriers to each relay, the scheme minimizes exposure to full key information. Furthermore, when a relay forwards Alice's channel information to Bob, the information is anonymized by using the reciprocal channel between the relay and Bob. This ensures that channel information is protected from potential eavesdroppers. Compared with conventional relay-assisted physical layer key-generation schemes, the proposed TL-PLKG significantly improves the secret key rate (SKR) by a factor of 14.14 times.

The primary contributions of this study are as follows:

1) A relay-assisted wireless communication system was proposed to improve the coverage scalability of next-generation Wi-Fi.

2) A physical layer key-generation algorithm, TL-PLKG, was introduced to enhance the security of relay-assisted Wi-Fi systems.

3) A link adaptation technique is proposed to optimize the performance of the TP-LKG algorithm based on the relay locations.

The remainder of this chapter is organized as follows. Section 3 introduces the system model and the proposed scheme. Section 4 presents

the simulation setup and the performance evaluation results of the proposed method. In Section 5, the findings are discussed and future research directions are outlined. Finally, Section 6 concludes the study.

TABLE 3.1
Overview of prior physical layer key generation approaches in relay-assisted scenarios.

Ref.	Scheme	Relay Type	Main Contribution	Limitation
Kong et al., 2018	PLKG in relay networks	Trusted	- Generates secret keys using channel impulse response - Enhances performance via adaptive	- Assumes trusted relay; cannot prevent key leakage via the relay
Aldaghri and Mahdavifar, 2020	PLKG with induced randomness in static environments	Untrusted	- Alice and Bob simultaneously transmit random symbols to the relay - Each node removes self-interference to extract the other's randomness	- Requires perfect time synchronization - Self-interference removal must be precise
Thai et al., 2016	Multi-relay PLKG with untrusted relays	Untrusted	- Multiple-antenna legitimate nodes in presence of multiple relays - Evaluates security under non-/partial-/fully	- Synchronization errors degrade performance - Accurate relay gain estimation is required

			-colluding relay assumptions - Extracts key from superimposed relay signals	
Keshavarzi et al., 2024	Lightweight PLKG for 6G networks	Untrusted	- Embeds random phases into probe signals - Secret key is derived by removing self-phase and recovering the peer's phase	- Sensitive to synchronization mismatches
Xu et al., 2024	PLKG using cooperative jamming	Untrusted	- Alice and Bob simultaneously send training sequences - Relay only observes composite channel, preventing channel separation	- Requires strict synchronization - Practical recovery from amplified signals is challenging

2. Related Work

Previous PLKG studies in relay-assisted environments typically assumed either trusted or untrusted relay settings (Table 3.1).

Kong et al. (2018) addressed the problem of key leakage by external eavesdroppers during the key-sharing process in conventional two-way relay networks. It generates a common key between two legitimate nodes based on the channel impulse response of multipath channels and enhances performance by dynamically adjusting the quantization parameters according to the signal-to-noise ratio (SNR) level through an adaptive quantization algorithm. The key-generation and quantization techniques in this study realize a secure network coding system that is robust against eavesdropping attacks. However, a limitation of this study is the assumption of a trusted relay, which makes it vulnerable to key leakage by the relay itself.

Aldaghri and Mahdavifar (2020), Thai et al. (2016), Keshavarzi et al. (2024), and Xu et al. (2024) have proposed key-generation techniques in untrusted relay environments, where the relay should not be able to access the secret key. These studies commonly adopt a strategy in which Alice and Bob simultaneously transmit signals to the relay, causing it to receive only a superimposed signal, thus preventing the extraction of individual channel information.

Aldaghri and Mahdavifar (2020) proposed an induced-randomness-based key-generation technique to overcome the low key-rate problem in static environments where wireless channel variation is minimal. In a relay-assisted setup, Alice and Bob simultaneously transmit random

symbol sequences to the relay to prevent both the relay and the eavesdropper from recovering individual randomness. By removing self-interference from the superimposed signal received from the relay, each user can recover a random sequence from the other and generate a final secret key.

Thai et al. (2016) considered a scenario with multiple untrusted relays and proposed a PLKG technique in which legitimate nodes with multiple antennas generate keys. The relays act according to a defined protocol while also serving as potential eavesdroppers. This study evaluated security under non-colluding, partially colluding, and fully colluding relay assumptions. To ensure secure key generation, Alice and Bob simultaneously transmit probing signals to the relays and reconstruct the counterpart's channel information by removing their own components from the amplified signals of the relay.

Keshavarzi et al. (2024) proposed a lightweight PLKG solution for 6G networks in untrusted relay environments. In this study, Alice and Bob independently select random phases, embed them in their probing signals, and later remove their own phase components to extract the randomness of the others, enabling secure key generation.

Xu et al. (2024) introduced a cooperative jamming technique for PLKG under untrusted relay settings. In this approach, Alice and Bob simultaneously transmit training sequences to the relay, forcing it to observe only the combined channel response, thereby preventing individual channel estimation. This study demonstrates that the proposed method effectively protects key confidentiality, even when Alice, Bob, Relay, and

Eve have spatially correlated channels.

However, a common limitation of the existing PLKG studies on untrusted relays is the requirement for perfect time synchronization between Alice and Bob. These studies assume ideal synchronization but do not propose practical methods to achieve it. Any mismatch in timing may significantly degrade key-generation performance. Additionally, because these studies often assume AF relays, accurate relay-gain estimation is crucial for reconstructing the counterpart's channel information and removing self-interference, which is another practical challenge.

In this study, we propose a multilink-based key-sharing mechanism that allows Alice and Bob to securely exchange channel information without requiring perfect time synchronization. Furthermore, we propose a DF-relay-based channel-sharing algorithm that enables key exchange without estimating relay gain or performing self-interference cancellation.

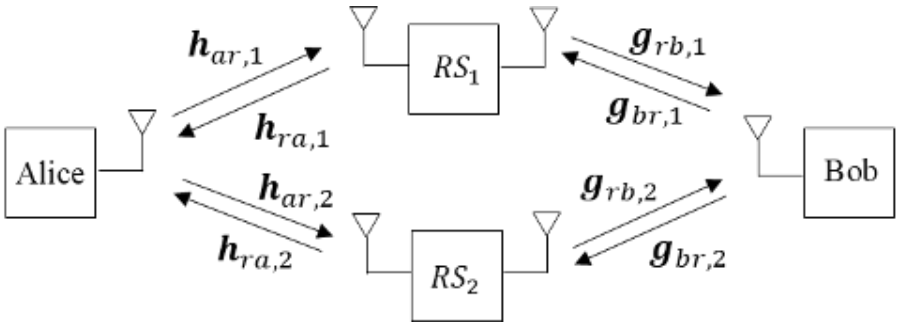

FIGURE 3.1 System model of TL-PLKG.

3. System Model and Proposed Method

1) System Model

Fig. 3.1 illustrates the system model of the proposed TL-PLKG scheme. In this setup, direct communication between Alice and Bob is unavailable; thus, they rely on the relay nodes to establish connections. Alice and Bob aim to generate a physical layer secret key using CSI to encrypt their messages while ensuring that the full key information is not exposed to relay stations (RSs).

The common randomness used for key generation is the CSI between Alice and the RSs. Bob securely obtains this CSI from the RSs without disclosing key-related information to them. Although RSs are assumed to operate according to the designated protocol (i.e., non-colluding), they are considered untrusted in the sense that they may attempt to infer the secret key.

Alice and Bob are multilink devices (MLDs) capable of multilink operations, and the two RSs are 2-way relays that can simultaneously communicate with both Alice and Bob. Furthermore, the RSs employ the

DF protocol, in which signals received from Alice or Bob are decoded and re-encoded before being forwarded.

The system model considered in this study consisted of two legitimate nodes, Alice and Bob, and two untrusted relays, RS_1 and RS_2. The wireless channel from Alice to the k-th relay, RS_k ($k \in \{1,2\}$), is denoted by $\boldsymbol{h}_{(ar,k)} \in \mathbb{C}^{(N \times 1)}$, where N represents the number of subcarriers in an orthogonal frequency division multiplexing (OFDM) symbol. The channel from RS_k to Alice is represented by $\boldsymbol{h}_{(ra,k)} \in \mathbb{C}^{(N \times 1)}$. The reciprocal channel between Alice and RS_k is referred to as $\boldsymbol{h}_k$. Similarly, the channel from Bob to RS_k is denoted by $\boldsymbol{g}_{(br,k)} \in \mathbb{C}^{(N \times 1)}$, and the channel from RS_k to Bob is denoted by $\boldsymbol{g}_{(rb,k)} \in \mathbb{C}^{(N \times 1)}$. The reciprocal channel between Bob and RS_k is denoted $\boldsymbol{g}_k$.

Before data transmission, Alice, RS_k, and Bob transmit only the preambles to probe the wireless channel and generate the physical layer secret keys. Both Alice and Bob are MLDs, and the communication between Alice and RS_k and between RS_k and Bob is conducted via a multilink operation, where the total bandwidth is evenly divided across k links. Each k-th link operates over a dedicated 10 MHz sub-band and is assigned $N = 26$ subcarriers out of the total 52 subcarriers in an OFDM symbol. The links are time-multiplexed within a single time slot and sequentially transmit data. In this system, Alice and Bob quantize the channel responses of the 52 subcarriers received over two links to generate a 52-bit secret key.

The channel pairs $(\boldsymbol{h}_{ar,1}, \boldsymbol{h}_{ar,2})$, $(\boldsymbol{h}_{ra,1}, \boldsymbol{h}_{ra,2})$, $(\boldsymbol{g}_{rb,1}, \boldsymbol{g}_{rb,2})$, and $(\boldsymbol{g}_{br,1}, \boldsymbol{g}_{br,2})$

were modeled as correlated fading channels, where the correlation was determined by the interrelay distance. The correlation coefficient is computed using the zeroth-order Bessel function of the first kind as $\rho = J_0(d_{RS1RS2}/d_c)$, where d_c denotes the correlation distance. d_c refers to the distance at which the channel correlation is reduced by approximately half. In this study, d_c was set to 20 m, and d_{RS1RS2} is the distance between the two relays. For example, channel $\boldsymbol{h}_{ar,2}$ can be expressed as

$$\boldsymbol{h}_{ar,2} = \rho \boldsymbol{h}_{ar,1} + \sqrt{(1-\rho^2)}\boldsymbol{w}. \tag{3.1}$$

where $\boldsymbol{w} \sim CN(\boldsymbol{0}, \boldsymbol{I})$. Similarly, the other channel pairs—($\boldsymbol{h}_{ra,1}$, $\boldsymbol{h}_{ra,2}$), ($\boldsymbol{g}_{rb,1}$, $\boldsymbol{g}_{rb,2}$), and ($\boldsymbol{g}_{br,1}$, $\boldsymbol{g}_{br,2}$)— were modeled in the same manner to reflect the spatial correlation between the two relays.

2) Performance Evaluation Metrics

The performance metrics used in this study to evaluate the key-generation scheme are the key mismatch rate (KMR), key leakage rate (KLR), and secret key rate (SKR).

1) KMR: The KMR refers to the probability that the generated keys will not match during the entire sampling process. For example, if key generation fails in 10 out of 100 sampling attempts, the KMR is 0.1. In this study, a key-generation attempt is considered successful only when all key bits are identical between legitimate parties.

2) KLR: The KLR refers to the probability that a given relay RS_k can fully reconstruct the secret key, resulting in a key identical to that generated by Alice or Bob. Because each relay receives only a subset of the OFDM symbol subcarriers, it inherently has access to only partial channel information. As the secret key is determined based on a full set of subcarriers, an individual relay cannot directly infer the entire key.

However, because the channels between the relays are spatially correlated, depending on their physical separation, the likelihood that RS_k can infer the channel values of the other link increases as the distance between the relays decreases. Although the relays are assumed to be non-colluding, the spatial correlation may allow a relay to estimate the channel values assigned to opposite links. Thus, KLR is defined as the probability that at least one relay can reconstruct the entire secret key owing to the spatial correlation with the opposite link.

3) SKR: The SKR, expressed in bits per second, quantifies the effective rate of secure key generation over time, considering both the KMR and

the KLR. In this study, the SKR is defined as shown in Eq. (3.2), where L denotes the length of the generated key (in bits) and t denotes the time required to complete one key-generation session.

$$SKR(bits/\text{sec}) = \frac{L \times (1-KMR) \times (1-KLR)}{t} \tag{3.2}$$

3) Channel Probing Procedure in TL-PLKG

Before generating a symmetric key based on reciprocal channel information, Alice and Bob must first probe the channel to estimate its value. Fig. 3.2 illustrates the procedure by which Alice, RS_k, and Bob perform channel probing over four timeslots.

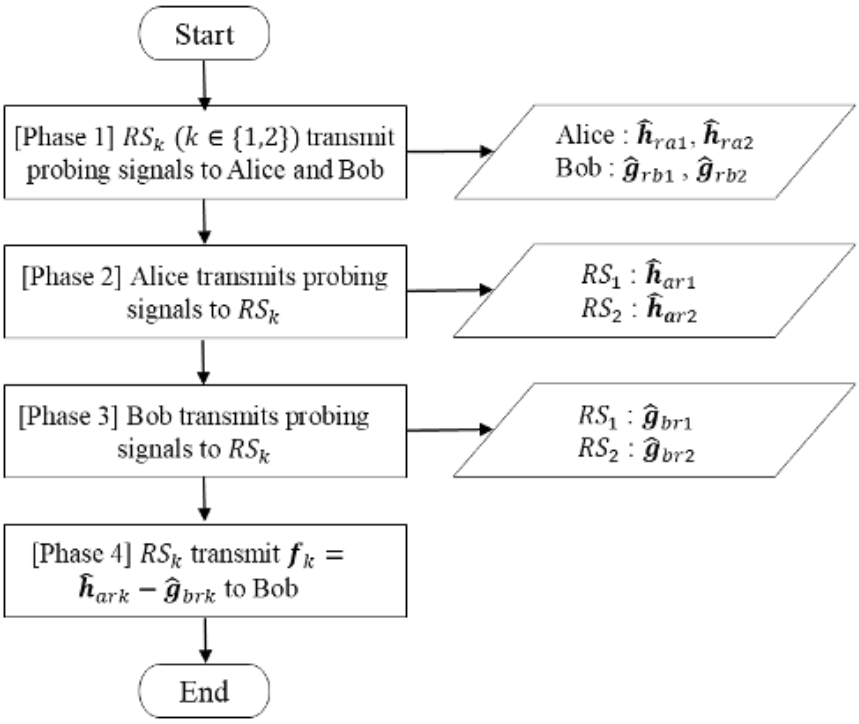

FIGURE 3.2 Flowchart of channel probing process.

In the first time slot, RS_1 and RS_2 transmit probing signals to Alice and Bob, respectively. Upon receiving the probing signals, Alice and Bob can respectively estimate the channels $\hat{h}_{ra,k} = h_{ra,k} + z_{ra,k}$ and $\hat{g}_{rb,k} = g_{rb,k} + z_{rb,k}$, where $z_{ra,k}$ and $z_{rb,k}$ denotes the channel-estimation errors. In the second

timeslot, Alice sends probing signals to RS_k in a time-division manner, allowing RS_k to estimate the channel $\hat{\boldsymbol{h}}_{ar,k} = \boldsymbol{h}_{ar,k} + \boldsymbol{z}_{ar,k}$. In the third timeslot, Bob similarly transmits probing signals to RS_k, enabling RS_k to estimate the channel $\hat{\boldsymbol{g}}_{br,k} = \boldsymbol{g}_{br,k} + \boldsymbol{z}_{br,k}$. In the final time slot, RS_k transmits the value $\boldsymbol{f}_k = \hat{\boldsymbol{h}}_{ar,k} - \hat{\boldsymbol{g}}_{br,k} \in \mathbb{C}^{(N\times 1)}$ to Bob.

The reason for transmitting $\boldsymbol{f}_k$ instead of $\hat{\boldsymbol{h}}_{ar,k}$ is that $\hat{\boldsymbol{h}}_{ar,k}$ is directly used in key generation, and transmitting it openly risks exposing channel information to another relay or external eavesdroppers. Because the channel between RS_k and Bob is also assumed to exhibit channel reciprocity, Bob can recover the estimated channel $\hat{\boldsymbol{h}}_{ar,k}$ by adding $\hat{\boldsymbol{g}}_{rb,k}$ to the received $\boldsymbol{f}_k$.

Finally, Bob generates the secret key based on the reconstructed channel $\tilde{\boldsymbol{h}}_{ar,k} = \hat{\boldsymbol{h}}_{ar,k} - \hat{\boldsymbol{g}}_{br,k} + \hat{\boldsymbol{g}}_{rb,k} = (\boldsymbol{h}_{ar,k} + \boldsymbol{z}_{ar,k}) - (\boldsymbol{g}_{br,k} + \boldsymbol{z}_{br,k}) + (\boldsymbol{g}_{rb,k} + \boldsymbol{z}_{rb,k})$, whereas Alice uses the locally estimated channel $\hat{\boldsymbol{h}}_{ra,k} = \boldsymbol{h}_{ra,k} + \boldsymbol{z}_{ra,k}$ for key generation.

4) Key-Generation Process in TL-PLKG

Alice and Bob generate the secret key by quantizing their respective estimated channel responses $\hat{\boldsymbol{h}}_{ra,k}$ and $\tilde{\boldsymbol{h}}_{ar,k}$. The quantization process compares each channel vector component to its average value; elements greater than the mean are quantized as 1, and the rest as 0.

Specifically, Alice computes the mean of her estimated channel vector $\hat{\boldsymbol{h}}_{ra,k}$, denoted as $\bar{h}_k^{(Alice)}$, and generates the binary key vector $\boldsymbol{K}_{Alice,k} \in \{0,1\}^{(\mathrm{N}\times 1)}$ by assigning 1 to elements larger than $\bar{h}_k^{(Alice)}$, and 0 otherwise.

Bob performs the same operation using his channel estimate $\tilde{\boldsymbol{h}}_{ar,k}$ and its mean $\bar{h}_k^{(Bob)}$, resulting in $\boldsymbol{K}_{Bob,k} \in 0,1^{(\mathrm{N}\times 1)}$. Finally, Alice and Bob concatenate the quantized key bits from all k channels to form the final secret keys, $\boldsymbol{K}_{Alice} = [\boldsymbol{K}_{Alice,1}; \boldsymbol{K}_{Alice,2}]$, $\boldsymbol{K}_{Bob} = [\boldsymbol{K}_{Bob,1}; \boldsymbol{K}_{Bob,2}]$.

5) Comparison of Local and Relay-aided Channel Estimation

The common randomness used by Alice and Bob for key generation is the reciprocal channel between Alice and RS_k, denoted by $\boldsymbol{h}_k$. However, Alice's key generation relies on locally estimated channels, whereas Bob's relies on reconstructed channels based on relay-assisted information. Consequently, the probability of a mismatch between the keys generated by Alice and Bob increases.

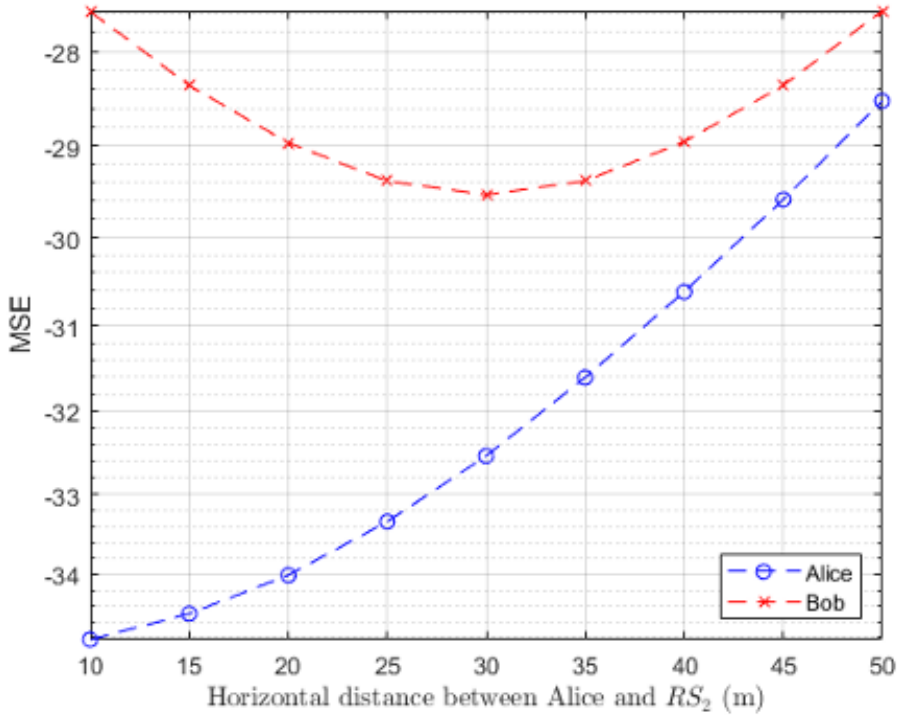

FIGURE 3.3. MSE of channel estimates at Alice and Bob with varying RS_2 position.

Fig. 3.3 shows the mean square error (MSE) between Alice's and Bob's channel estimates in a scenario where Alice and Bob are 60 m apart, RS_1 is placed at the midpoint between them, and RS_2 moves from a location close to Alice to a location close to Bob. In this setup, the SNR between Alice and RS_1 is approximately 26 dB.

The MSE was evaluated between the perfect channel h_k and the channel estimated by Alice, as well as between $\boldsymbol{h}_k$ and the channel reconstructed by Bob. Because Alice directly estimates her own channels with the relays, the MSE increases as the relay moves farther away. For Bob, the reconstructed channel vector is derived from the channels estimated by the relay and Bob himself; thus, the MSE is lowest when RS_2 is at the midpoint and increases as RS_2 moves closer to either Alice or Bob. Moreover, because Bob's reconstructed channel contains both the relay's estimation error and his own, Bob's MSE is consistently higher than Alice's MSE.

6) Key Leakage Scenario

RS_k can acquire the channel information of its own link during the channel-estimation phase and thus obtain a partial view of the overall key. However, it cannot directly observe the channel information associated with the subcarriers assigned to other relays and must rely on eavesdropping to infer the remaining part.

For instance, when two relays are deployed, each relay is responsible for 26 subcarriers, allowing it to acquire 26 key bits under ideal channel conditions. The remaining 26 bits can be inferred by overhearing the channel used in another relay. The spatial correlation between the relays is modeled as shown in Eq. (3.1), where a shorter distance between relays increases the probability of accurately estimating the channel of the other relay, thereby recovering the full key.

If at least one relay successfully reconstructs a key that matches the one generated by Alice and Bob, it is considered a leakage event, and the KLR is used to quantify this risk.

4. Performance Evaluation

1) Experiment Environment

In this study, we consider a two-dimensional coordinate system in which legitimate nodes and relays are placed as follows: Alice is located at $(x_{A,}y_A)=(0,0)$ and Bob is located at $(x_{B,}y_B)=(x_{b,0})$. Two relay nodes, RS_1 and RS_2, were placed symmetrically with respect to the x-axis. RS_1 is located at $(x_{R1,}y_{R1})=(x_b/2,\lceil x_b/3 \rceil)$, and RS_2 is located at $(x_{R2,}y_{R2})=(x_b/2,-\lceil x_b/3 \rceil)$.

A log-distance path loss model was employed with path loss exponent $n=3.5$, carrier frequency f=2.4 GHz, and speed of light c=3×108 m/s. The reference path loss at distance $d_0=1$ m is calculated as $PL_{d_0}=10\log_{10}(\frac{(4\pi f)^2}{c^2})$.

The distances between Alice and RS_1, RS_2 are denoted as d_{h1} and d_{h2}, respectively, and the distances between Bob and RS_1, RS_2 are denoted as d_{g1} and d_{g2}. Distances were calculated using the Euclidean distance formula. The path loss for each link is then given by $PL_{h1}=PL_{d_0}+10n\log_{10}(\frac{d_{h1}}{d_0})$, and similarly for PL_{h2}, PL_{g1}, and PL_{g2}.

Given the transmit power P_t, the noise power spectral density N_0, and the bandwidth B, the noise power is given by $N_0 \cdot B$, and the received power in dB for each link is $P_r=10\log_{10}(P_t)-PL$, and the SNR in dB is computed as $SNR=P_r-10\log_{10}(N_0 \cdot B)$. The transmit power P_t of Alice, Bob, and each RS_k is uniformly set to 0.1 W.

In the experiments, the performance was evaluated under varying SNR conditions between Alice and RS , ranging from 0 to 36 dB. This corresponds to a decrease in the distance between Alice and Bob from 300 to 30 m. Therefore, this study assessed the performance and security of the proposed method for Alice-Bob distances ranging from 300 m to 30 m.

To evaluate the performance of the proposed TL-PLKG, we compared it with two representative relay-based PLKG schemes from previous studies (Kong et al., 2018; Thai et al., 2016). Kong et al. (2018) assumes a single trusted AF relay and generates secret keys based on the channel responses that traverse the relay link (trusted-relay scheme).

By contrast, Thai et al. (2016) considers a scenario involving multiple untrusted AF relays, where Alice and Bob simultaneously transmit signals to the relays. The relays then forward the superimposed signals back, enabling Alice and Bob to extract each other's channel information from the aggregated observations (superposition-based multirelay scheme).

For a fair comparison with TL-PLKG, all three methods were evaluated under the same frequency band, bandwidth, and transmission power. In the case of superposition-based multirelay scheme, we assumed a multilink environment with two relays, which is consistent with the setup used in this study.

2) Key Mismatch Rate

Fig. 3.4 compares the KMR of the proposed TL-PLKG with prior relay-based approaches. As the distance between Alice and Bob decreased, the received SNR increased, resulting in a lower KMR for all schemes. The trusted-relay scheme shows the lowest KMR because it directly quantizes the signals received from the relay for key generation, thereby minimizing errors. In contrast, the superposition-based multirelay scheme suffers from higher mismatch rates because of the need to estimate the relay gain and extract the counterpart's channel values from the superimposed signals.

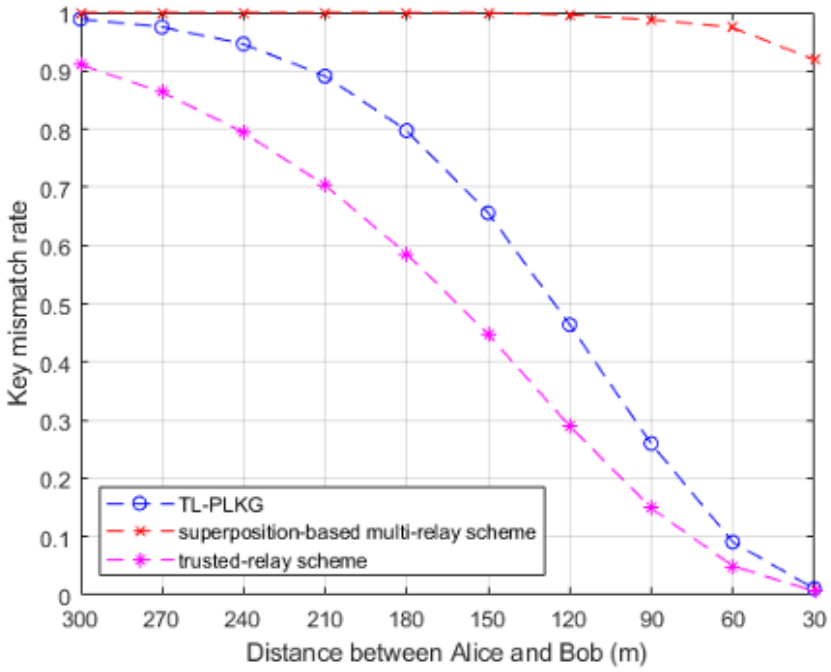

FIGURE 3.4 KMR performance of TL-PLKG vs. Prior relay-based approaches.

Moreover, both the TL-PLKG and trusted-relay schemes generate 52-bit keys, whereas the superposition-based multirelay scheme generates a 208-bit key according to the proposed method of Thai et al. (2016). Consequently, the probability that all 208 bits match is naturally lower than that of a 52-bit key, which leads to the higher KMR observed in Fig. 3.4. This highlights a trade-off: longer key lengths are advantageous from a key leakage perspective but increase the probability of a key mismatch.

3) Key Leakage Rate

To ensure a fair comparison with prior studies, we assume that, as in our proposed scheme, all relay nodes in the baseline schemes can acquire channels between Alice and Bob through channel probing via malicious behavior. In multirelay environments, the likelihood of obtaining complete channel information increases as the distance between relays decreases. In single-relay environments, the closer the relay is to Alice and Bob, the more accurately the relay can estimate the channels, thereby increasing the probability of generating a key identical to the legitimate key generated by Alice and Bob.

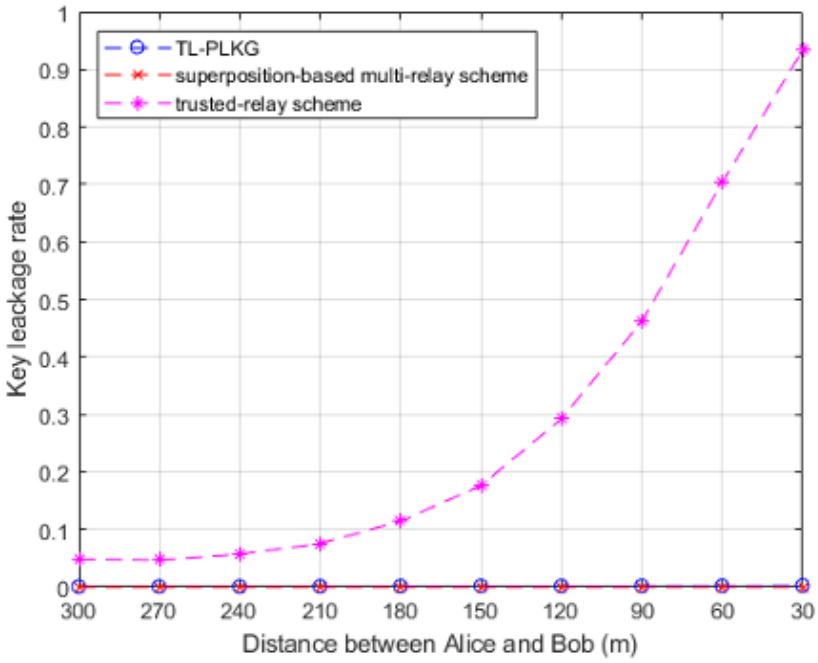

FIGURE 3.5 KLR performance of TL-PLKG vs. Prior relay-based approaches.

Fig. 3.5 compares the KLR of the proposed TL-PLKG with those of the existing schemes. In the trusted-relay scheme with a single relay, KLR increases sharply as the distance between Alice and Bob decreases. In contrast, both the TL-PLKG and superposition-based multirelay schemes achieved nearly zero KLR across the tested range.

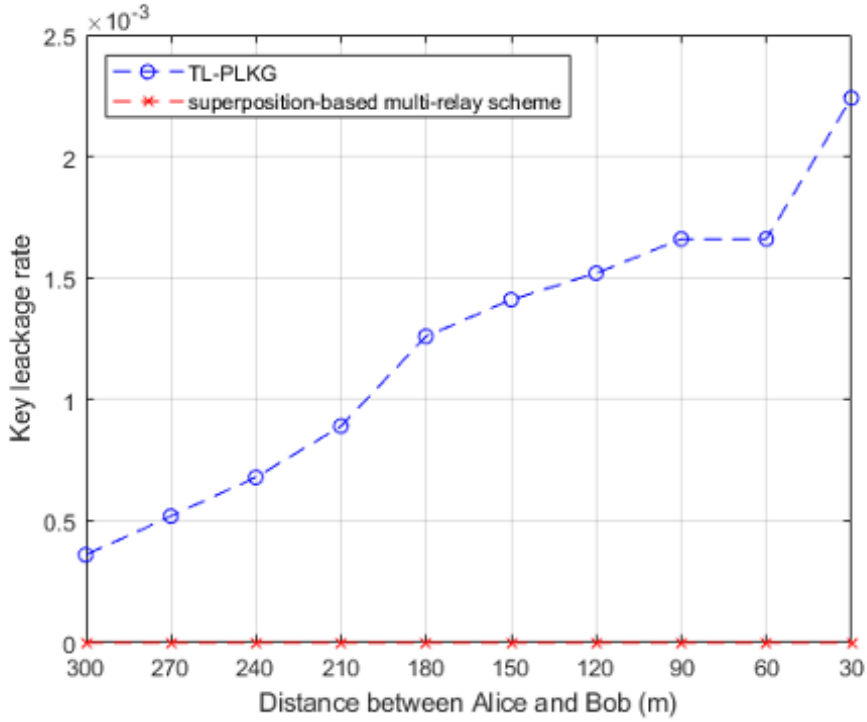

FIGURE 3.6 KLR performance of TL-PLKG vs. superposition-based multirelay scheme.

Fig. 3.6 compares the KLR between the TL-PLKG and the superposition-based multirelay scheme. Although TL-PLKG generates a 52-bit key, the superposition-based scheme generates a 208-bit key, which makes it more robust from the perspective of key leakage.

4) Secret Key Rate

To compare the key-generation performance of the proposed scheme and existing methods, we evaluated the key-generation speed in terms of the SKR. As defined in Eq. (3.2), SKR (bits/s) represents the length of the key that can be securely generated per unit time without leakage. The key-generation time $t = 16us \times \tau$ refers to the time required to transmit a $16us$ preamble during the entire time slot τ.

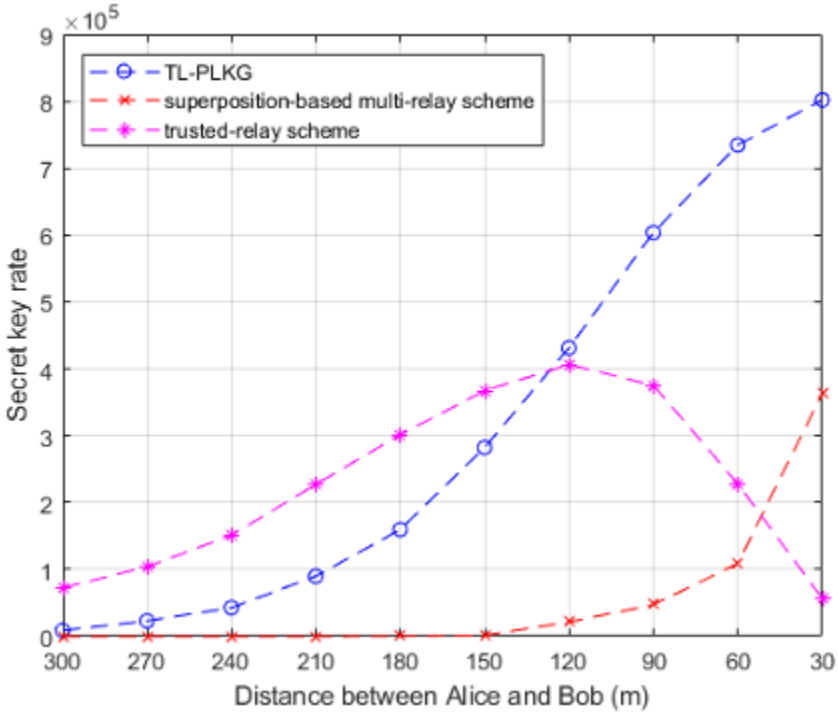

FIGURE 3.7 SKR performance of TL-PLKG vs. Prior relay-based approaches.

In TL-PLKG and the trusted-relay scheme, a 52-bit key is generated over 4 time slots ($L = 52$, $\tau = 4$), while in the superposition-based scheme,

a 208-bit key is generated over 3 time slots ($L=208$, $\tau=3$). The SKR reflects the probability of successfully generating an L-bit key over time t, without mismatch or leakage.

Fig. 3.7 compares the SKR of the proposed TL-PLKG scheme with that of conventional relay-based PLKG schemes, illustrating the upper bound of the SKR performance. Although the superposition-based scheme generates keys four times longer and uses fewer time slots, it suffers from a high KMR, resulting in the lowest SKR among the three schemes. The trusted-relay scheme achieves the highest SKR at distances beyond 120 m as KLR gradually decreases. However, its SKR significantly drops at shorter distances owing to a sharp increase in the KLR.

In contrast, the proposed TL-PLKG demonstrated the highest SKR in the 30-120 m range, where the SNR conditions were favorable. At 30 m, TL-PLKG achieves an SKR that is 2.21 times higher than that of the superposition-based scheme and 14.14 times higher than that of the trusted-relay scheme.

Conventional relay-based key-generation methods exhibit a trade-off within 120 m: either high key-generation performance comes with a high leakage risk, or low leakage comes at the cost of reduced key-generation. However, the proposed TL-PLKG addresses this challenge by minimizing both KMR and KLR, thereby achieving robust key-generation performance even within short to medium distances.

5. Discussion

In this study, we proposed a TL-PLKG scheme in a non-colluding untrusted relay environment. However, if the relays share the channel information, the probability of reconstructing the entire key increases significantly. In future work, we aim to develop a secure key-generation protocol in the presence of colluding, untrusted relays. For instance, relay selection algorithms can be devised to allow Alice and Bob to choose a subset of relays for key generation among multiple untrusted relays, thereby mitigating the risks posed by collusion. Moreover, the randomization of the subcarrier indices used for key generation can be explored to prevent further information leakage.

6. Conclusion

This chapter proposes the TL-PLKG algorithm, which enables secure secret key generation between two legitimate users in the presence of untrusted relays without exposing the relays to key-related information. In the TL-PLKG scheme, each relay is assigned only a portion of the overall OFDM subcarriers, thereby preventing any single relay from accessing the full set of channel information used for key generation. Furthermore, when relaying the forward channel information, reciprocal channels are used to anonymize the data, thereby mitigating the risk of channel information leakage to eavesdroppers or other untrusted relays. While conventional PLKG schemes suffer from either a high KLR or high KMR in medium- to long-range scenarios, the proposed TL-PLKG achieves up to 14.14 times higher SKR than prior methods when the

distance between the two legitimate users is within 120 m. In future work, we aim to investigate secure key-generation schemes that remain robust even in environments with colluding relays.

Table of Notations (Chapter III)

Notations	Description
RS	Relay station
a	Alice
b	Bob
k	Index of relay stations
$h,\ g$	Channel
N	The number of subcarriers in an OFDM symbol
ρ	The channel correlation coefficient
d	Distance
z	The channel estimation error
K	The quantized secret key
x	x-axis of Alice, Bob, and relay stations
PL	Path loss
P_t	Transmit power
P_r	Received power
L	Key length
τ	The number of time slot

Ⅳ. Secure Digital Identity Management System

1. Introduction

The digitalization of information systems has advanced rapidly, and service structures are progressing toward Web 3.0, which emphasizes users' complete control over their data, as characterized by de-centralization, openness, and data ownership (Gan et al., 2023). Considering the main objective of improving data ownership and individual privacy (Sambra et al., 2016), technologies and policies focusing on data ownership were introduced in various domains, including finance, medical data, and electronic notifications. MyData is a primary service in the finance sector that strengthens data sovereignty by allowing individuals to manage their financial data. Techniques for authenticating and identifying data subjects ensure data ownership and utilization (Satybaldy et al., 2024; Perugini et al., 2024). In the financial sector, secure authentication and identification (ID) methods are critical owing to the increasing prevalence of online electronic financial transactions such as fintech and non-face-to-face financial payments (Khan et al., 2023; Meng et al., 2019).

Digital transformation, led by governments, is rapidly advancing across various industries, including the financial sector, and e-government, which promotes digital transformation, has become a global trend (Aubert and Chan, 2024; Senyo et al., 2024). Traditional ID technologies, such as identification cards and passports, are also becoming digitalized. Likewise, digital ID serves as a core function in managing information within the

information systems of governments, private sectors, and administrative institutions. Estonia launched the e-Estonia project in 2002, thereby establishing an e-government system that connects various sectors such as healthcare, banking, education, and law online. The key technology underlying X-Road, which serves as the backbone of e-Estonia, is universal electronic identification, with digital signatures granted to all citizens at the age of 15. Through a reliable national identity scheme, Estonia successfully developed an electronic government (e-government), and the trust of citizens in the integrity of the government system was a critical factor in its successful adoption (Anthes, 2015).

Five countries, namely, the United Kingdom (UK), South Korea, Israel, New Zealand, and Estonia, formed Digital 5 (D5) and are in collaboration for the development of digital governments. To construct an e-government system in which all services are integrated online, it is essential to consolidate the information of all citizens into a government database. The UK attempted to introduce a national identity system, which failed owing to public concerns over privacy and liberties (Anthes, 2015).

Attempts to establish national identity systems for the introduction of e-government are a current global trend. The European Union (EU) introduced digital identity (eID) as part of its policy to secure digital sovereignty in an open and interconnected world and to realize a human-centered, sustainable, and prosperous digital future. In particular, the objective of eID is to provide all key public services online by introducing a digital ID that stores personal information in electronic form.

Finland drafted the National ID Program legislation in 2022 and piloted a digital ID network in which citizens could store credentials on their preferred digital wallet. Similar to the Estonian X-Road concept, India aimed to build an “identity database” for all citizens by collecting biometric data, including photographs, fingerprints, and iris scans.

The resident registration number is a national personal ID number issued to all citizens living in South Korea. It is “unique” because all citizens are issued different numbers according to their date of birth, gender, and place of birth. Once issued, the number exhibits “invariance” as it can only be re-issued under certain conditions. Although the national identity system has the advantage of maximizing convenience and efficiency by allowing for the prompt ID and online management of all citizens, several privacy concerns exist. The response of the Republic of Korea to the COVID-19 pandemic is considered exemplary (Moradi et al. 2020; Kang et al. 2020). During the COVID-19 pandemic, the tracing system in Korea, which was developed by collecting location data, credit/debit/prepaid transaction data, closed circuit television (CCTV) footage, immigration records, transit data, and personal identification data, enabled social network analyses such as the identification and blocking of the spread of infected people, the identification of super-spreaders, and the detection of group infection cases (Kim et al., 2021). The inclusion of the personal ID information of individual citizens as a collectible element in the case of an emergency simplified the tracking of infected people by the Korea Centers for Disease Control and Prevention (Lee and Choi, 2020). A personal identity tracking system using resident registration numbers was

effective in crises across the country; however, there is a significant privacy problem. Identifying and releasing the path of activities containing personal private information have associated issues (Lee and Choi, 2020), thereby giving rise to privacy problems in a system that manages individuals by providing them with fixed identifiers (Jeon et al., 2025; Lee et al., 2025).

Digital identity authentication in South Korea was initially dependent on the resident registration number of individuals. However, more personal information than required was collected to verify the identity of an individual, and policies to limit the collection of resident registration numbers were implemented to solve this issue (Kim et al., 2018). Thus, alternative means of identification are required. The MyData industry uses connecting information (CI) to identify individuals. CI technology replaces resident registration numbers as a means of digital ID and is generated by the one-way encryption of the resident registration number, thereby rendering it unique and unrecoverable. Its application to various daily life services such as identity authentication, asset integration, tax payments, and electronic notifications has increased user convenience and reduced social costs.

CI is secure because it uses a hash algorithm that is cryptographically difficult to restore to plain text. However, it exhibits the same vulnerability as traditional resident registration numbers because one-way encryption with the same resident registration number and secret key combination always generates the same CI. Thus, the resident registration number and CI can be matched individually, thereby enabling user

tracking. Tracking the one-way encryption algorithm employed to generate the CI for the original message is challenging. Thus, a resident registration number based solely on the CI cannot be generated. However, the secret information used to generate the CI is generated, stored, and managed by the identity verification organization and certificate authority. The same CI is generated and provided using the same information whenever a CI generation request is made. Moreover, CI is used, disused, and discarded according to the privacy policy based on the Personal Information Protection Act. However, the identity verification service provider may retain and use it until the outsourcing contract is terminated. From CI creation to the transmission, processing, and storage of service data utilizing CI, the leakage of CI due to system and network vulnerabilities may result in the leakage of users' personal information. However, there has been little research on analyzing the vulnerabilities arising from the immutability of CI (Palamà et al., 2021).

CI uses simple cryptographic techniques to identify individuals across multiple organizations while using the domestic identity authentication framework based on the resident registration number. Therefore, CI has been used in the domestic digital identity utilization industry for identity verification and identification through regulatory sandboxes. Moreover, CI was legislated as a technology for identifying individuals across different organizations in the future. However, to securely share and utilize data in the digital identity industry, conventional CI technologies should be complemented by vulnerability analyses of the utilized services. Therefore, the vulnerabilities arising from the uniqueness and immutability of CI

must be analyzed, and secure methods for its utilization must be explored. The vulnerabilities of CI generation and utilization technologies were analyzed in this study, and a secure digital identity technology to alleviate CI vulnerability was developed. The proposed digital identity mechanism generates multiple digital identities according to requirements and periodically updates the digital identity values. Thus, even if one identity is leaked, the other identities cannot be inferred. Moreover, because the same identity is consistently updated, information leakage is minimized.

The contributions of this study are as follows:

1) The technical structure and features of CI for the identification of individuals across different organizations were analyzed, in addition to security vulnerabilities and threats at the CI generation stage.

2) The technical structure and operating principles of the financial sector (MyData), in addition to mobile electronic notification services that utilize CI, were analyzed. Furthermore, CI-related security vulnerabilities and attack scenarios that may arise during service provision and data transmission were investigated.

3) Finally, security-enhanced CI technology for the safe use and activation of digital identity-enabled services was proposed, and the performance and security levels of conventional CI and the proposed method were evaluated.

The remainder of this chapter is organized as follows. Section 2 presents an analysis of prior research on CI utilization services. Section 3 presents an examination of the vulnerabilities of CI technology by analyzing the process of CI creation and utilization. Section 4 proposes

secure CI technology to improve CI vulnerabilities. Section 5 presents the discussion and future research directions, and Section 6 concludes the study.

2. Related Work

1) Digital Identity Management Systems

TABLE 4.1
Previous studies of digital identity management systems.

Ref.	Year	Parameters	Considerations
Yu Y et al.	2016	Cloud, Cryptography	Computation cost, Communication cost, Storage cost
Norrman et al.	2016	Cryptography	Session-based pseudonym update for 5G privacy
Saeed et al.	2021	Cryptography	Lightweight pseudonym mutable IMSI privacy protection
Yang and Li	2020	Blockchain	Time cost, Cost, Throughput
Liu et al.	2020	Blockchain	Authentication, Privacy, Trust
Stockburger et al.	2021	Decentralized ID	SSI, Transparency, Interoperability
Samir et al.	2021	Decentralized ID	Secret sharing, Smart contract, SSI, Reliability

Recent research on digital identity systems was focused on preventing identity leakage during the transmission and storage of digital identities. In particular, several studies were focused on decentralized identity technologies using blockchain technology. Table 4.1 lists previous studies focused on privacy issues related to digital identity.

Yu et al. (2016) proposed a method for constructing an identity-sharing system without revealing the identity of plaintext based on cryptographic technology in a cloud environment. They improved the complexity and cost of remote data integrity check (RDIC) protocols to ensure secure data storage in cloud computing environments. The ID-based RDIC protocol achieved ZKPs without disclosing the stored data to the verifier. Despite improvements in the efficiency and complexity of data protection

systems, there were vulnerabilities associated with the fuzzy vault scheme when combining biometric information with keys, and the system complexity increased owing to the presence of a third-party auditor in a cloud environment.

Privacy concerns related to the traceability of digital identities have already been recognized as a critical issue in the field of mobile communication technologies. The international mobile subscriber identity (IMSI) is a unique identifier used when a mobile device connects to a network, and attackers can exploit this to perform IMSI catching attacks to track users or determine their locations. Norrman et al. (2016) and Saeed et al. (2021) proposed a countermeasure to IMSI catching by replacing IMSI with a pseudonym, which is an encrypted version of the IMSI. In both studies, the pseudonym is generated by the user equipment (UE) and the home network using a public key or session key and is used instead of the IMSI to prevent its exposure. Furthermore, a new pseudonym is generated and updated each time the session changes, thereby eliminating the linkability and traceability of the IMSI. However, while the digital identifiers in these studies are used for authentication in a one-to-one relationship between the UE and the home network, applying such mechanisms to online services would require multiple service providers to simultaneously update and synchronize the pseudonym to identify the user consistently. The technique proposed in this study differs in that it not only supports user authentication but also enables multiple online service providers to continuously update and synchronize the CI, thereby allowing secure and privacy-preserving user

identification in a multi-service environment.

Yang and Li (2020) and Liu et al. (2020) identified issues in centralized identity systems and proposed a blockchain-based identity management system. Yang and Li (2020) used a blockchain to alleviate the problems of centralized identity systems because traditional centralized digital identity management systems are subject to threats such as single-point failures, internal attacks, and privacy leakage. Moreover, the lack of privacy confidentiality due to the inherent transparency of blockchains was enhanced by modifying the existing claim identification model of the blockchain using smart contracts and ZKPs to realize unidentifiability and limit the exposure of the ownership of attributes. However, blockchain-based systems pose a risk of identity information exposure or abuse. Additional measures can, therefore, be implemented to increase the complexity and reduce the efficiency of identity management systems.

Stockburger et al. (2021) applied a blockchain-enabled decentralized identity management system to the field of public transportation. Traditional identity management systems are prone to data breaches and identity theft, and the public transportation sector, in particular, requires integrated identity management across multiple operators and countries. A self-sovereign identity (SSI) system, which allows users to manage and control their own identities, reduces reliance on central authorities while enhancing user privacy and security. SSI-based identity systems can promote security, transparency, and interoperability across systems. Samir et al. (2021) proposed an SSI-based digital identity management service that combines secret sharing techniques with smart contract technology.

The framework stores the identity credentials of internet of things (IoTs) in external storage to prevent tampering or misuse. This framework can play a critical role in enabling a trustworthy decentralized identity management system in IoT environments, supporting secure service exchanges, data collection, and decision-making among IoT devices.

Previous studies were focused on secure data storage, transmission, and authentication in identity management systems; however, privacy issues stemming from the uniqueness and invariability of digital identities remain unsolved. Consequently, a security-enhanced digital identity system is required to mitigate existing privacy and complexity issues.

2) Digital Identifier-based Identity Management Systems

TABLE 4.2
Previous studies of digital identifier-based identity management systems.

Ref.	Year	Identifier Technology	Main contributions
Bossenko et al.	2024	Unique Identifier	e-ID-based digital patient identification system
Ramamoorthi et al.	2024		The implementation of Federated Digital Identifiers (FDIs) in healthcare systems
Manoj et al.	2022	DID-based Identifier	Applying blockchain-based DID technology for user authentication in Electronic Health Records (EHR) systems
Waleed et al.	2023		Enhancing trust, security, and interoperability in IoT systems based on DID technology
Yin et al.	2022		Strengthening security for identity authentication and data exchange among IoT devices and ensuring privacy protection through the blockchain-based DID system, SmartDID.

Systems utilizing digital identifiers across various fields have been proposed to identify the identities of people and objects. Digital identifiers are particularly in high demand in healthcare and medical sectors, where accurate and non-redundant differentiation of entities is crucial for recording and utilizing data. Furthermore, in recent years, decentralized identity (DID) technology has gained attention for its ability to distinguish network nodes in a decentralized manner without central intervention,

especially in networks like IoT.

Bossenko et al. (2024) designed a patient identifier system based on Estonia's e-ID and X-Road platforms. This study proposed a system that assigns a unique identification number to individuals using e-ID and connects all medical records to it. Additionally, encryption technology and access control through user privilege management were applied to ensure the protection of patient data. The findings demonstrated that Estonia's digital patient identification system has the potential to enhance healthcare data management and patient-centered services. Ramamoorthi et al. (2024) reviewed the implementation status, related technologies, challenges, and future prospects of Federated Digital Identifiers (FDIs) in healthcare systems. FDIs are a digital identification framework that enables user authentication and data sharing across different institutions and systems. With a single digital identifier, users can access multiple organizations, and data are managed in a decentralized environment. The study highlights that continued development of FDIs requires technological integration, standardization, and policy support.

Recent studies have proposed decentralized identity management systems that utilize DIDs as identifiers. Manoj et al. (2022) introduced a blockchain-based DID technology to be applied for user authentication in Electronic Health Records (EHR) systems. Traditional EHR systems rely on centralized databases, making them highly vulnerable to data breaches and hacking. Moreover, they often lack robust authentication mechanisms to establish trust among hospitals, healthcare providers, and patients, and patients frequently cannot fully control their own data. By leveraging a

DID-based EHR system, users can directly manage their identity information and data, while the blockchain foundation ensures immutability and transparency. Waleed et al. (2023) investigated the current technological capabilities and potential applications of DIDs in IoT systems. DIDs have the potential to revolutionize trust, security, and interoperability among IoT devices, enhancing the autonomy and efficiency of IoT networks. However, the study emphasizes the need to address challenges such as scalability and the lack of standardization, while focusing on the development of energy-efficient and security-centered DID solutions. Yin et al. (2022) proposed SmartDID, a blockchain-based DID system designed to ensure privacy protection and trust in IoT environments. SmartDID enhances the security of identity authentication and data exchange between IoT devices while safeguarding user privacy. IoT devices generate DIDs and register verifiable credentials on the blockchain, utilizing smart contracts to perform DID-based authentication. Additionally, privacy-enhancing technologies such as ZKPs and encryption are employed. However, the study emphasizes the need to address the limitations of blockchain network performance, resource constraints in IoT devices, and the lack of standardization.

While several studies have explored identity authentication and identity management technologies, they have primarily focused on aspects such as data stability and privacy, highlighting a significant limitation. Specifically, there is a lack of research addressing issues arising from the use of unique identifiers. This study focuses on analyzing the limitations of identifier technologies and proposes security measures for identifiers.

3. Analysis of the Vulnerabilities of Digital Identity Technologies

1) Analysis of the Structural Characteristics and Vulnerabilites of CI

① CI generation and issuance process

CI refers to 88 bytes of information generated by an identity verification organization using the one-way encryption of the resident registration number and is used online to identify the same user across different Internet service providers. Eq. (4.1) is the CI generation equation. The terminology for the information elements of CI generation is presented in Table 4.3.

$$CI = HMAC_{sk}((RN \| Padding) \oplus S_{A)} \tag{4.1}$$

TABLE 4.3
Definitions of the terms for the CI generation equation.

Information	Description
HMAC	Keyed-Hash Message Authentication Code (SHA 512)
RN	Resident registration number (13 bytes = 104 bits)
Padding	408 bits are filled with “0x00 00 ... 00”, with the exception of 104 bits for the resident registration number to make the input value 512 bits.
S_A	Secret information shared among identity verification organizations (64 bytes = 512 bits)
sk	Secret key shared among identity verification organizations (64 bytes = 512 bits)

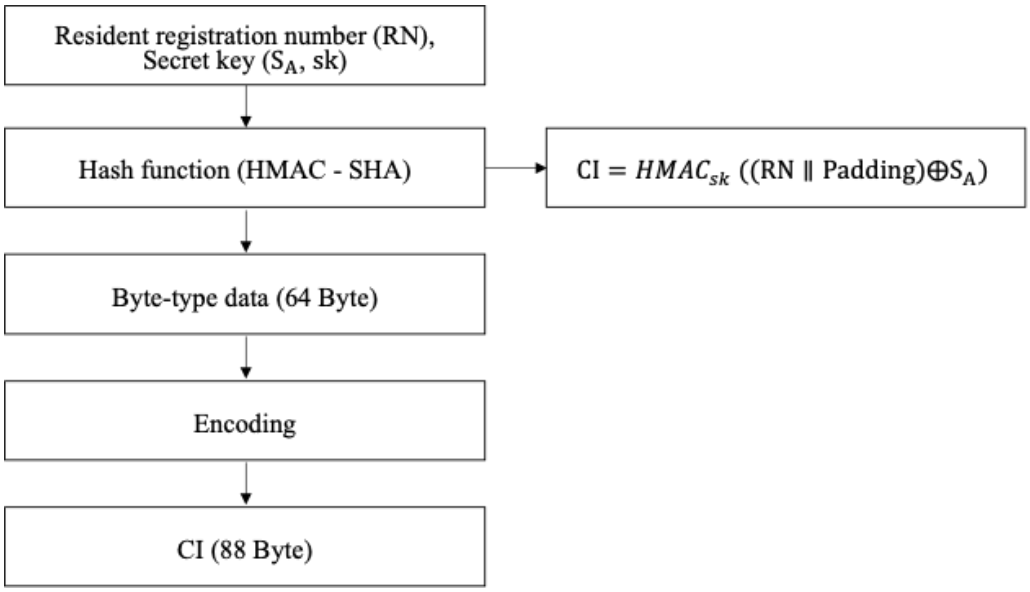

FIGURE 4.1 Connecting information generation flowchart.

Fig. 4.1 presents a flowchart of the CI generation process. The first input for generating the CI is the resident registration number (104-bit unique identification number comprising 13 unique digits). The SHA 512 hash algorithm used to generate the CI requires 512 bits of input. The identity verification organization adds 408 bits of zero padding to the user resident registration number (*RN*) to generate duplicate subscription verification information. Subsequently, the generated value is exclusively XORed with 512-bit secret information (S_A) shared among identity verification organizations, and it is encrypted with a 512-bit secret key (*sk*) shared among identity verification organizations. A hash-based message authentication code (*HMAC*) is calculated to generate a 64-byte hash value (64 bytes) and is processed into a CI of 88 bytes by encoding.

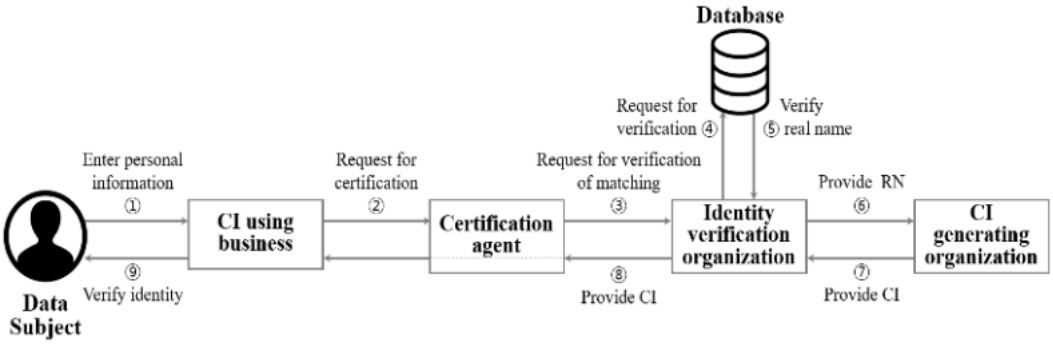

FIGURE 4.2 Connecting information issuance process.

Fig. 4.2 presents the process of issuing CI between a CI-using business, certification agent, and identity verification organization. First, when the data subjects subscribe to a service that uses CI:

(1) They provide personal information to CI-using businesses to verify their identities.

(2) The CI-using business that receives personal information requests the certification agent to authenticate the identity of the data subject.

(3) The certification agent requests that the identity verification organization verifies the identity of the data subject.

(4) The identity verification organization requests verification from the database containing the personal information of the data subject (e.g., the mobile phone subscription database).

(5) The database returns real-name verification results to the identity verification organization.

When the personal information entered by the data subject matches the database information, the identity verification organization provides the

generating organization with the resident registration number of the data subject.

(6) The CI-generating organization generates a CI via the above process using the resident registration number of the data subject and delivers it to the identity verification organization.

(7) The generated CI is provided to the CI-using businesses via each organization. The CI provided to the CI-using business should be encrypted and stored securely.

(8) Finally, the CI-using business confirms the identity of the data subject, and the process is terminated.

② Analysis of CI vulnerabilities

CI encrypts resident registration numbers using a strong hash algorithm (SHA 512). Consequently, restoring the resident registration number using backward operations is technically difficult. However, an analysis of the CI generation principle indicates that CI is generated using a fixed resident registration number and a secret key (S_A, sk) generated by an identity verification organization. Thus, the same value is generated each time. Furthermore, although the CI should be generated as a unique value, the same value could be generated depending on the secret key S_A, which is the input $(RN \| Padding) \oplus S_A$ to the hash-based message authentication code ($HMAC$) for generating the CI. Although individual resident registration numbers (RN) are unique and are intended to guarantee the uniqueness of the CI, collisions may occur due to certain values of the secret key S_A, resulting in identical $HMAC$ inputs.

Moreover, during the CI issuance process, personal information may be leaked during the identity authentication process (①-⑤) or CI information may be leaked during the process of providing the generated CI (⑦-⑧). Fig. 4.3 depicts the probability of CI theft by a man-in-the-middle (MITM) attack while issuing the CI. The CI provided by the CI-generating organization is delivered to the CI-using business through an identity verification organization and certification agent. Consequently, the CI can be leaked through a MITM attack owing to transport layer security (TLS) and API vulnerabilities. A CI can be readily generated if the information required to generate the CI is leaked while generating, issuing, or utilizing CIs. Moreover, the indiscriminate use of CI can lead

to the leakage of personal information by tracking and specifying individuals.

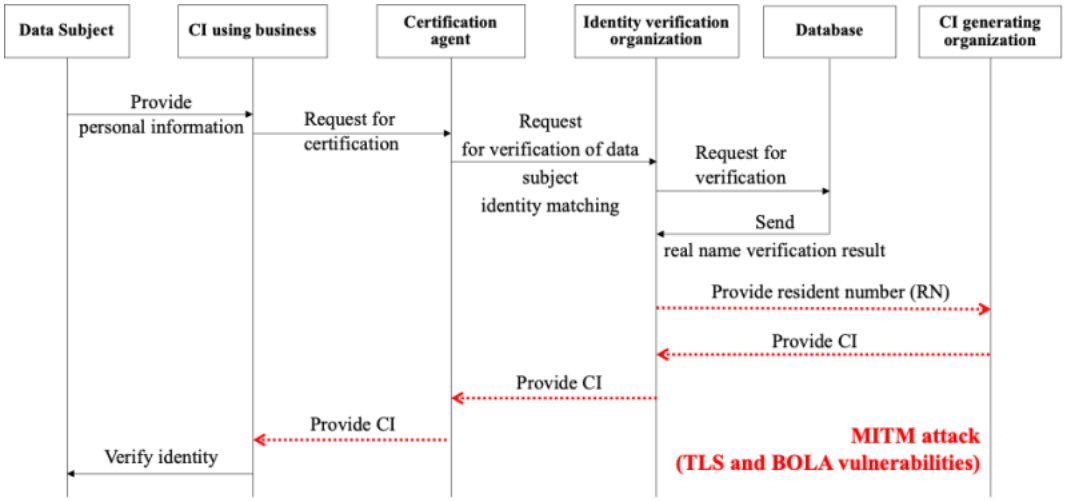

FIGURE 4.3 Vulnerabilities in the connecting information issuance process.

2) Analysis of the Structure and Vulnerabilites of CI Utilization Technologies

CI is utilized in various services (MyData and mobile electronic notification services) to strengthen the data economy and establish data sovereignty. This section presents an analysis of the technical characteristics of representative services that utilize CI, and a discussion on the security vulnerabilities and threats related to CI leakage that may occur when utilizing these services.

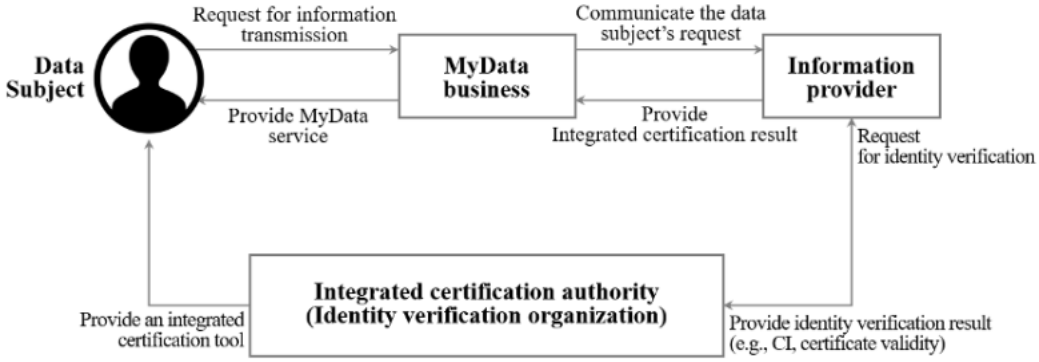

FIGURE 4.4 Integrated authentication process for the MyData service.

① Financial MyData service

Fig. 4.4 depicts the integrated authentication process of the MyData service (National Radio Research Agency, 2012). The MyData service comprises four objects: the customer as a data subject, the MyData business, the information provider, and the integrated certificate authority. The MyData business refers to a provider who comprehensively collects

data distributed to the data subject and provides various services. The information provider stores and manages the data of the data subject and provides data when the data subject is requested by the information provider. The integrated certificate authority is an identity verification organization that verifies user identity when the MyData service requests access to user data. When a customer intending to use the MyData service requests the MyData business to send information, it delivers the information requested by the customer to the information provider. Subsequently, the information provider requests that the data subject authenticate its identity. After authentication, the information provider transmits the information requested by the customer to the MyData business. The data transmission process is terminated when the information provider delivers this information to the customer. During the transmission of information within the MyData service, integrated certification authorities and information providers utilize CI to identify customers.

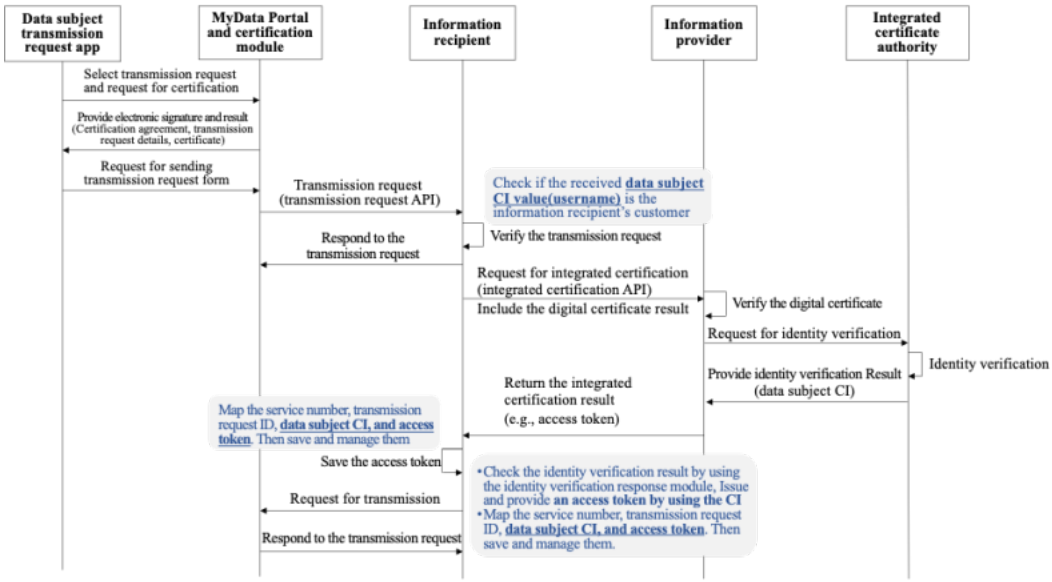

FIGURE 4.5 Access token issuance process for asset list inquiry.

The MyData business and information provider is required to follow the pre-defined standard API transmission specifications and procedures to transmit and receive personal credit information of the data subject between them. To request the transmission of personal credit information using the API, the recipient of the information receives an access token to confirm its eligibility to access the personal credit information of the customer held by the information provider. Fig. 4.5 presents the access token issuing process via the MyData Portal after the identity is verified via the integrated certificate authority to view the list of assets held in the MyData service. To request assets from the information provider, the data subject generates authentication information such as an electronic signature using an integrated authentication method. The MyData Portal requests the issuance of an access token to the information provider

using an integrated authentication API. The information provider then sets permissions based on the transmission request history of the data subject and requests an identity verification process to the certification server of the integrated certificate authority. The integrated certificate authority provides the identification results, including the CI of the data subject, to the information provider, which then issues an access token to the MyData portal to finalize the integrated authentication. Thereafter, the information recipient receives the data from the information provider (as pertaining to the data subjects) using the access token. CI is used to verify the identity of the data subject when issuing access tokens for using MyData API. The CI and access tokens are mapped and then stored and managed. Consequently, an unauthorized entity could use a vulnerable CI to obtain an access token to acquire assets and gain access to data from the data subjects.

② Mobile electronic notification service

The mobile electronic notification service delivers electronic bills to the user's smartphone through a certified electronic document intermediary instead of conventional paper delivery. Bills and notices contain sensitive and personal information, which involves the risk of personal information being leaked when mailed. The mobile electronic notification service reduces the probability of personal information leakage via the conventional method as it requires legitimate identity verification prior to viewing the notice, and the contents are transmitted in an encrypted manner.

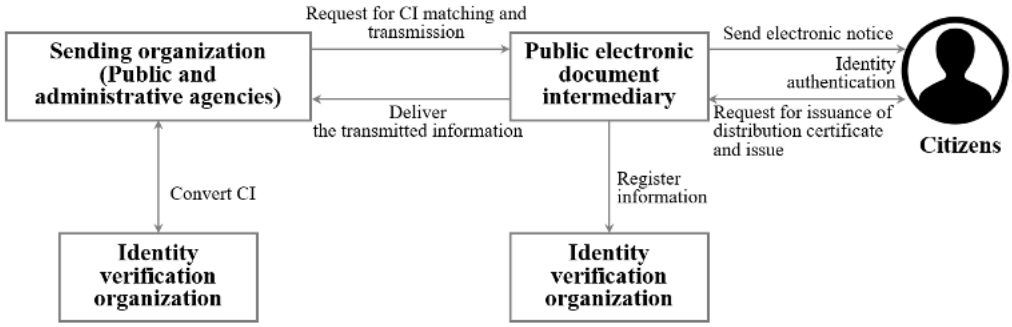

FIGURE 4.6 Integrated authentication process in the mobile electronic notification service.

Fig. 4.6 presents the integrated authentication process in the mobile electronic notification service, which comprises the sending organization, identity verification organization, certified electronic document intermediary, electronic document distribution center, and recipient of the notification. The sending organization is a public or administrative organization that provides the user resident registration number to the

identity verification organization for the provision of mobile electronic notification services. The identity verification organization converts the resident registration number into a CI and forwards it to the sending organization, which delivers the CI received and the notice to be received by the user to the certified electronic document intermediary. Thereafter, notices are sent to users matched to each CI. Users pass through an identity verification process and read the notices. The Electronic Document Distribution Center registers and stores the distribution information of notices and manages the public electronic addresses used to issue distribution certificates. Thus, public and administrative organizations, in addition to electronic document distributors, use CI for user identification in mobile electronic notification services. However, personal and sensitive information may be leaked because users can be tracked through CI identification values that are commonly used by multiple organizations.

③ Security threats to services utilizing CI

CI is employed to identify individuals in the services provided by public, administrative, and private organizations. However, if the CI is leaked during the data transmission process of a service due to the structural vulnerability of CI technology, it may have the same impact as the leakage of unique identification numbers. The CI values generated by an identity verification organization are the same values for all organizations, and are not updated. If CIs are compromised, users may be unaware of the breach and become vulnerable to targeted attacks on their personal information. This section presents an analysis of the security threats related to CI information leakage that may occur in the data transmission process of MyData and electronic notifications. Table 4.4 presents the CI-related security threats that can occur in services that utilize CI.

TABLE 4.4
Security threats to services utilizing CI.

Security threats		Description	CI vulnerabilites
MITM	TLS vulnera bilities	- TLS 1.3 has compatibility issues with conventional protocols. - Issue pertains to backward compatibility. If TLS 1.3 is not supported, the system reverts to using TLS 1.2.	- When CI is used in MyData and electronic notification services, the CI may be leaked due to the transmission of personal information using an insecure transport protocol due to the TLS backward compatibility issue.

	BOLA	- Compromised object-level authentication. - If a compromised API is used, attackers can access and manipulate unauthorized objects. - Attackers can manipulate the identity of objects sent in an API request.	- The integrated certificate authority for MyData and electronic notification services requests CI from the information provider backbone system through the API, and the backbone system sends CI information in response to the request. - The occurrence of a broken object level authentication attack in the above process may lead to the theft of CI and users' personal information managed by the server and the integrated certificate authority.
CSRF		- Exploits the web server trust of requests from authenticated users. - Attacker manipulates tampered requests as legitimate requests through the browser of the pre-authenticated user.	- Unauthorized users can hijack access tokens mapped and managed with CI to steal personal and sensitive information of the data subjects.

1) MITM

An MITM involves the deceit of both participants in a communication system, to intercept or manipulate information. In the MyData service data transmission scenario, the MyData business and information provider transfer data while sharing the customer CI to identify the customer and verify their information. Thus, an attacker disguised as a MyData business or information provider can obtain customer data containing CI by intercepting packets during the data transmission process. Furthermore, when the information provider requests authentication from the integrated certificate authority in response to the customer identification request, the integrated certificate authority provides CI, certificate validity, and identity verification information. In this case, attackers can collect data in transmission using an MITM attack to gain information regarding the CI and individuals (Alwazzeh et al., 2020).

MyData uses the TLS protocol to protect the communication section. However, TLS vulnerabilities can lead to CI theft through MITM attacks (Lee et al., 2021). Transport layer security is a network security protocol that encrypts data for communication by encrypting a transmission section. It provides a secure connection that protects the privacy of two connected users and ensures data integrity. TLS 1.3 is recommended in the MyData industry because it is faster and more reliable than the previous versions. However, TLS 1.3 is subject to compatibility issues with existing protocols, which can be attributed to backward compatibility. If one party does not support TLS 1.3, the system reverts to using TLS 1.2 (Lee et al., 2020). Thus, if either the sending or receiving node does

not support TLS 1.3, TLS 1.2 can cause problems.

Broken object level authentication (BOLA) is among the 10 API security threats (2023) identified by the Open Web Application Security Project (OWASP) (OWASP, 2023a). If a web application uses an API with compromised object-level authentication, an attacker can access and manipulate unauthorized objects. Attackers can manipulate the identity of the objects sent in API requests, thereby resulting in data leakage, modification, loss, or account hacking. The integrated certificate authority requests the CI from the information provider backbone system through the API, and the backbone system sends the CI information. Therefore, if the abovementioned attack occurs in the MyData and mobile electronic notification services, CI theft and the theft of users' personal information managed by the service server and the integrated certificate authority may occur.

2) Cross site request forgery (CSRF)

In CSRF, an attacker exploits a web server's trust in requests from authorized users by sending a tampered request to the web server through the browser of a pre-authenticated user to the web server. This causes it to appear as a legitimate request (OWASP, 2023b). The MyData portal issues an access token from the information provider to the asset requested by the data subject. The integrated certificate authority verifies the identity of the data subject by verifying the electronic signature certificate and delivers the CI to the information provider. After authenticating the data subject using the CI, the information provider issues an access token for calling the MyData information provision API and sends it to the MyData portal. A CSRF attack primarily utilizes tokens generated by the server for manipulation attacks. An unauthorized user without access rights can obtain the access token and access the data of the data subject, thereby resulting in personal information leakage.

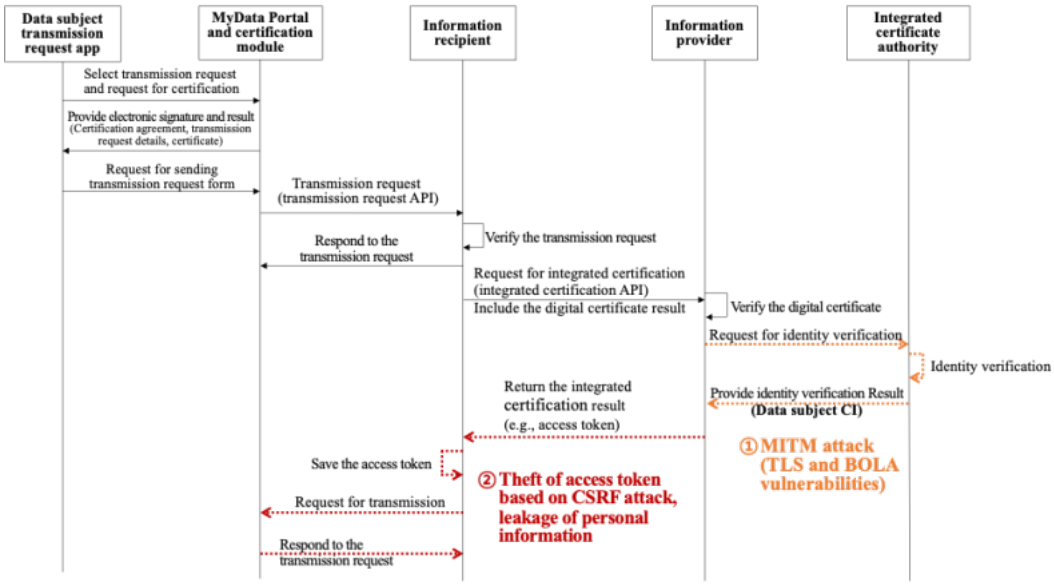

FIGURE 4.7 Possible attack scenarios during MyData access token issuance process.

④ Attack Scenarios

Fig. 4.7 presents the attacks that may occur during the access token issuance process based on the CI security threats. MITM attacks, shown in Fig. 4.7 ①, such as attacks based on TLS vulnerability or BOLA, can occur when an integrated certificate authority delivers the CI of a data subject to an information provider. First, a downgrade attack is launched from TLS 1.3 provided by the MyData service to TLS 1.2 by manipulating protocol messages to steal information during the communication process. The attacker manipulates the “hello” message sent from the client to the server as the client and server handshake communicate. Notably, the TLS 1.2 protocol is vulnerable to eavesdropping and tampering; thus, it can be used to obtain information

about the CI of a data subject by hijacking data during communication. Second, when the CI is provided using the API from an integrated certificate authority to an information provider, an attacker can steal information using TLS or BOLA attacks. A compromised API can be used in the information provision process to steal data by manipulating the IDs of objects contained in the API. Furthermore, attacks such as CI acquisition, account hacking, and data modification can be launched, resulting in the theft of the personal information of users managed by the integrated certificate authority.

When an integrated certificate authority delivers a CI to the information provider, the latter uses the CI to authenticate the data subject and issue an access token for API calls. If the access token is stolen via a CSRF attack during storage by the information recipient (② in Fig. 4.7), an unauthorized user can view the data of the subject using the token.

4. Digital Identity Technology with Enhanced Security

1) Proposed Technology

Conventional CI systems are used to identify individuals across different organizations, and each organization or service uses the same CI as the data source. This is the simplest and most efficient method in terms of system performance and user experience. However, the use of a single fixed CI is vulnerable to a single-point-of-failure problem, which may result in severe privacy issues. Therefore, a digital identity technology that can improve security while maintaining system performance is required. This study proposes an enhanced CI (eCI) system that generates multiple CIs from a single unique identification number and updates them periodically.

① System architecture

The CI concatenates 104 bits and 408 bits of resident registration numbers with zero padding, respectively, to obtain 512 bits (Eq. (4.1)). It is then XORed with a 512-bit secret key S_A. Subsequently, the generated value is encrypted with the secret key sk using the HMAC algorithm. To generate a unique CI, the input value to the hash algorithm should be unique for each data subject. When a resident registration number is XORed with a secret key, collisions may result in the uniqueness of the resident registration number being lost and the generation of the same value. CI uses two secret keys S_A and sk in contrast to duplication information (DI) (only one secret key sk). This increases the computational complexity and secret key management costs. The proposed

eCI generation system periodically updates the CI based on a one-time password (OTP) to preserve the uniqueness of the resident registration number and improve the inefficiency of the secret key.

The proposed eCI system first generates encrypted *RN* (*Temp*) data, which is expressed as follows:

$$Temp = HMAC_{sk}(RN \| SI) \tag{4.2}$$

where $CI = H(Temp \| PRN)$, *RN* is the resident registration number of the data subject, *PRN* is a pseudo-random number shared between the CI generating organization and the CI-using business, and the service identifier (*SI*) is the identification number of the CI-using business. Moreover, *sk* is a secret key shared by the CI generating organization. The CI generating organization concatenates the *RN* and *SI*, encrypts it using the secret key, and delivers the *RN* to the CI-using business.

A *PRN* can be generated using the time-based OTP technique, which is expressed as follows:

$$PRN = truncate(HMAC(K, C_T)) \bmod 10^d \tag{4.3}$$

The seed *K* for *PRN* generation is delivered by the CI generating organization. Subsequently, using the time of the CI generating organization authentication server and the CI-using operator OTP system, which are time synchronized, a counter value, $C_T = \lfloor \frac{T_C - T_0}{T_I} \rfloor$, is used ($T_0$ is the Unix time that starts a time-interval count with the default value can set to zero, T_I is a time interval used to calculate the counter value, and T_C indicates the current time). The CI update interval can be set to T_I. Eq. (4.3) is used to generate a PRN with d digits, and the CI

is updated in cycles of T_I.

The final equation for CI generation is expressed as follows:

$$CI = H(HMAC_{sk}(RN||SI)||truncate(HMAC(K, C_T)) \bmod 10^d) \quad (4.4)$$

The CI generating organization and CI-using business share a CI generation module. The encrypted RN (e.g., $Temp$ derived from Eq. (4.2)) value generated by the CI generating organization is entered as a fixed value into the CI generation module. Subsequently, all organizations that manage the CI update their own CI at regular intervals by generating a PRN using the seed K received from the CI generating organization and the OTP system counter value C_T of the CI generation module.

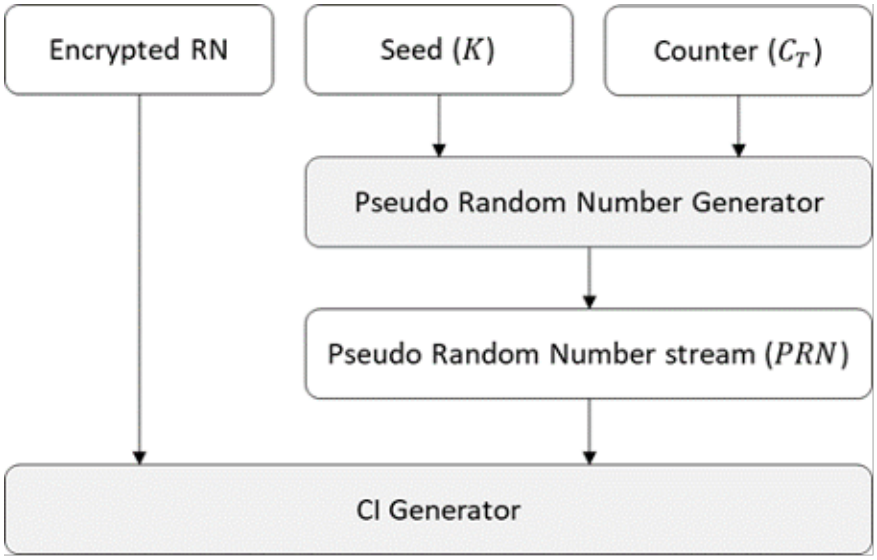

FIGURE 4.8 Structure of the connecting information generation module of the embedded connecting information system.

Fig. 4.8 presents the structure of the CI generation module. The CI generating organization concatenates the unique identification value of the data subject, *RN*, with the business identification number, *SI*; encrypts it to generate an encrypted *RN*, and delivers it to the CI-using business. The CI generation module comprises a time-based counter generator, a pseudo-random number generator, and a CI generator. The counter value is generated according to the cycle and inputted to the pseudo-random number generator with seed *K*, which outputs the *PRN*. The *PRN* and encrypted *RN* are operated through the CI generator and output an updated CI.

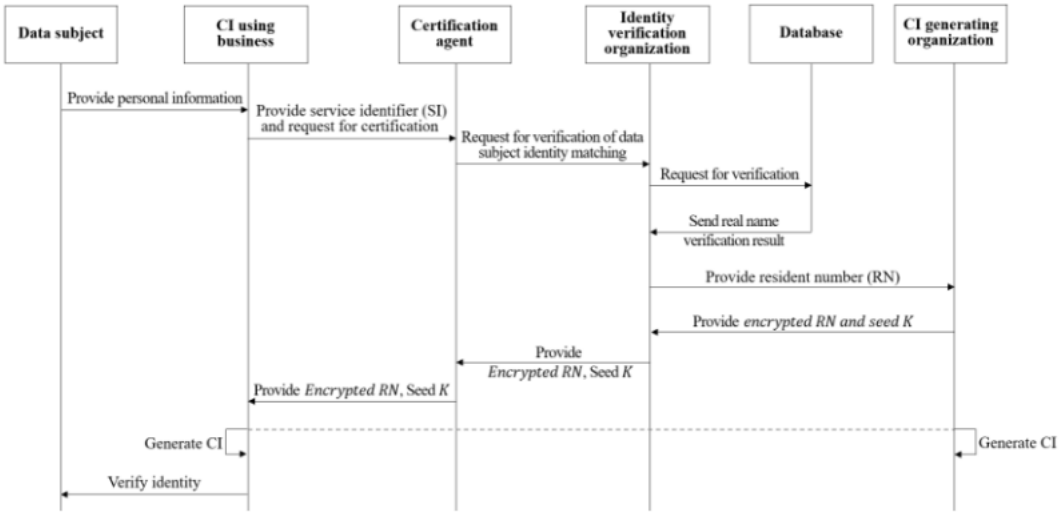

FIGURE 4.9 Connecting information issuance process of the enhanced connecting information system.

② Operation principle and utilization scenario

Fig. 4.9 presents the eCI issuing process. The CI-using business receives personal information from the data subject to generate CI and provides the personal information of the data-subject and the business identification number *SI* to the certification agent. The certification agent prompts an identity verification organization to verify the identity of the data subject, and the latter verifies the actual name through a certification database and provides the *RN* to the CI generating organization. The CI generating organization generates encrypted *RN*, *SI*, and seed *K*, which are delivered to the CI-using business through the identity verification organization and the certification agent. The CI generating organization and CI-using business use temporally synchronized OTP-based random number generators to generate the same CI and complete the identity

verification process. Consequently, CI-using and CI-creating organizations update their CIs at regular intervals through the OTP system. In summary, in CI-using businesses, only the encrypted *RN*—generated by combining the CI and *SI*—from the CI-generating organization, along with the seed value used for updates, is issued. The CI is then updated, synchronized over time, and utilized accordingly.

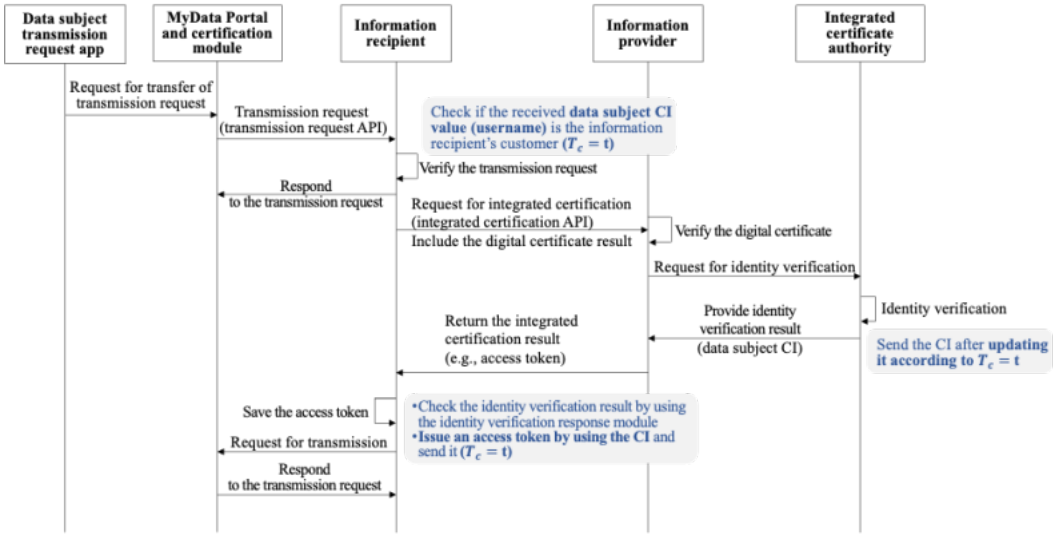

FIGURE 4.10 MyData access token issuance process based on the enhanced connecting information system.

Using the eCI system in the access token issuance process illustrated in Fig. 4.10, the information recipient and the integrated certificate authority update the CI at a fixed time interval to authenticate the identity and provide the access token. The information recipient updates the pre-issued CI periodically. Moreover, the integrated certificate authority stores the seed and time information when the CI is initially generated and provides the updated CI during the identity verification process. Thus, the

information recipient and the integrated certificate authority can use the updated CI to verify the identity of the data subject.

2) Security Evaluation

CI is used to identify the same user in the core services of society. Information from the same user of different services can be linked and combined owing to CI leakage, and significant amounts of personal information may be leaked. This section presents an evaluation of the impact of information leakage due to CI leakage to assess the security of the conventional and proposed methods.

① System model

In the conventional CI system, the impact of information leakage (I_{CI}) caused by CI leaked during the service period K is expressed as follows:

$$I_{CI} = \sum_{t=1}^{K} (w_t \times u \times p(1-p)^{t-1}), \tag{4.5}$$

where u is the number of data subjects (equal to the number of CIs in the conventional CI system), $p(0 \leq p \leq 1)$ is the probability of CI leakage in one transaction, and t implies transaction. For CI update cycle i and number of CI updates j, $K = i \times j$. In the conventional CI system, $j = 1$ and $K = i$. The impact of information leakage of the entire CI leak is obtained by multiplying the time weight w_t of CI leakage over the entire service period, the amount of CI information u, and the probability of CI leakage at a time point t for the entire transaction period. In conventional CI systems, the CI leakage time weight w_t ($w_t = K-t+1$) is calculated as the number of transactions from the time of CI leakage to the time of final service.

In the eCI system, a data subject CI is generated differently for each

service. Thus, the number of CIs increases proportionally with the number of services s to a total of $u \times s$; however, the amount of CI reduces by u/s. The impact of information leakage of the CI in an eCI system is the same as that in a CI system. Using the CI leakage time weight, the amount of CI, and the CI leakage probability at time t for the service period K in the eCI system, the impact can be expressed as follows:

$$I_{eCI} = \sum_{t=1}^{i \times j} (w_t \times \frac{u}{s} \times p(1-p)^{t-1}), \tag{4.6}$$

Unlike the conventional CI system, the eCI system was designed to reset the weight w_t in each round, where $w_t = r_t \times i - t + 1$, because the CI leakage effect is initialized whenever the CI is updated. Here, $r_t = \lceil \frac{t}{i} \rceil$ is the CI round at time point t, which is increased by one whenever the CI is updated.

To evaluate the performance of the eCI system in environments where CI is actually utilized, we constructed three representative scenarios—Government system, bank, and large-scale online service—and compared the impact of information leakage across these scenarios. $q(0 \le q \le 1)$ denotes the proportion of services—among the total services s used by a user—that have been compromised by an attacker or have experienced CI leakage. $\alpha(0 \le \alpha \le 1)$ represents the synchronization failure rate of the eCI system. The impact of information leakage in a real-world environments can be expressed as follows:

$$I_{eCI}^{real} = \sum_{t=1}^{i \times j} (w_t \times \frac{u \times q}{s} \times (1-\alpha) \times p(1-p)^{t-1}), \tag{4.7}$$

② Performance evaluation results

Table 4.5 presents the parameters used for performance evaluation. The numbers of data subjects u and of transactions during the service period K were set to 100 each. The total number of services s was set to increase by one, two, and four to model the generation of multiple CIs for data subjects according to s. For the conventional CI system, i is the same as K. The eCI system was set to update every 100, 50, and 25 transactions; thus, the update cycle i was shortened and the number of updates j increased. Further, p (probability of CI leakage) was set to increase from 0.01 to 0.1.

TABLE 4.5
Parameters for performance evaluation.

Parameter	Description	Value
u	Number of users	100
K	Number of transactions during the service period	100
s	Number of services	[1, 2, 4]
i	CI update cycle	[100, 50, 25]
j	Number of CI updates	[1, 2, 4]
p	CI leakage probability	[0.01: 0.01: 0.1]

Fig. 4.11 presents the impact of information leakage as a function of the number of services s used by the data subject and the number of CI updates as the probability of CI leakage p increased. In the conventional CI system, regardless of the number of services, the number of CIs of the data subject was the same as u, and the impact of information leakage of the conventional CI system was not influenced by s and j

because CI was not updated. In contrast, when the number of services (s) increased in the eCI system, the amount of CI information influenced by a single leakage decreased to u/s, which reduced the impact of information leakage because CI was separated by the objective of the data subject. Furthermore, shorter CI update intervals reduced the duration of CI leakage, thereby reducing the impact of information leakage due to CI leakage. Thus, the eCI system reduces the impact of CI leakage by distributing the CI data across multiple systems and shortening the CI update cycle.

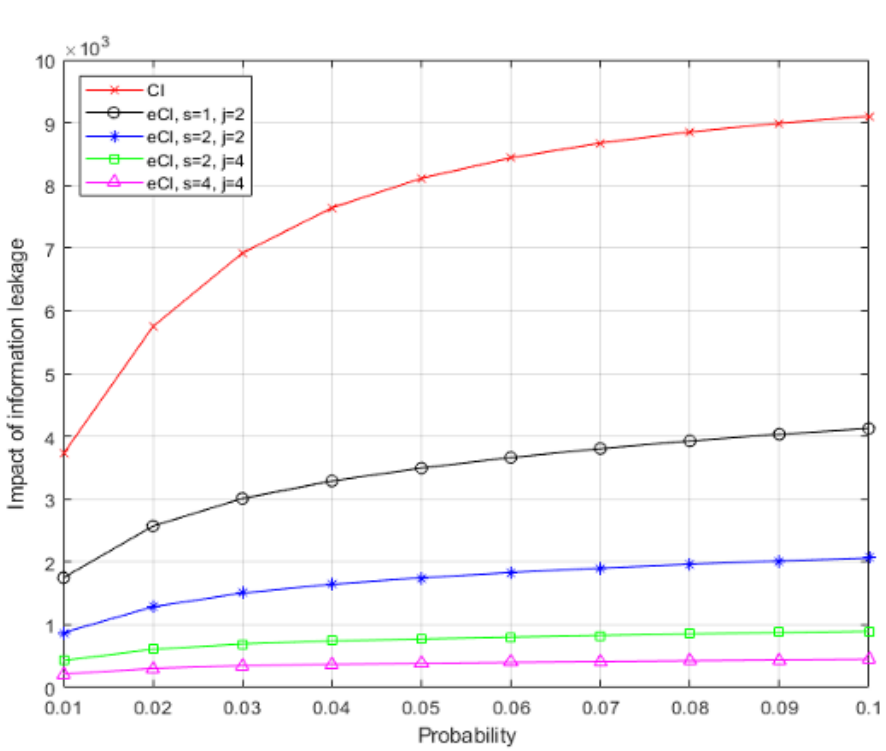

FIGURE 4.11 Connecting information leakage probability versus impact of information leakage.

TABLE 4.6
Parameters for performance evaluation of scenario-based simulation.

Parameter	Description	Value		
		Scenario A (Government)	Scenario B (Bank)	Scenario C (Large-scale online service)
u	Number of users	2,000	5,000	20,000
K	Number of transactions during the service period	100	100	100
s	Number of services	2	10	20
i	CI update cycle	50	50	50
j	Number of CI updates	2	2	2
p	CI leakage probability	[0.01: 0.01: 0.1]	[0.01: 0.01: 0.1]	[0.01: 0.01: 0.1]
q	The proportion of compromised services	0.2	0.5	0.8
α	The synchronization failure rate	0.05	0.1	0.2

Table 4.6 presents the parameters used for performance evaluation in the scenario-based simulation. Three representative cases with different levels of security and service scale were modeled: Scenario A represents a government system, Scenario B represents a banking environment, and Scenario C corresponds to a large-scale online service. The number of users was set to 2,000 for Scenario A, 5,000 for Scenario B, and 20,000

for Scenario C. The number of services used per user was configured as 2, 10, and 20 for each scenario, respectively. The CI update interval and total number of transactions were fixed across all scenarios. The value of q, which represents the proportion of services compromised by the attacker, was set to 0.2, 0.5, and 0.8, respectively. The parameter α denotes the synchronization failure rate of the system, and from the attacker's perspective, a higher synchronization failure reduces the success rate of attacks using leaked CIs.

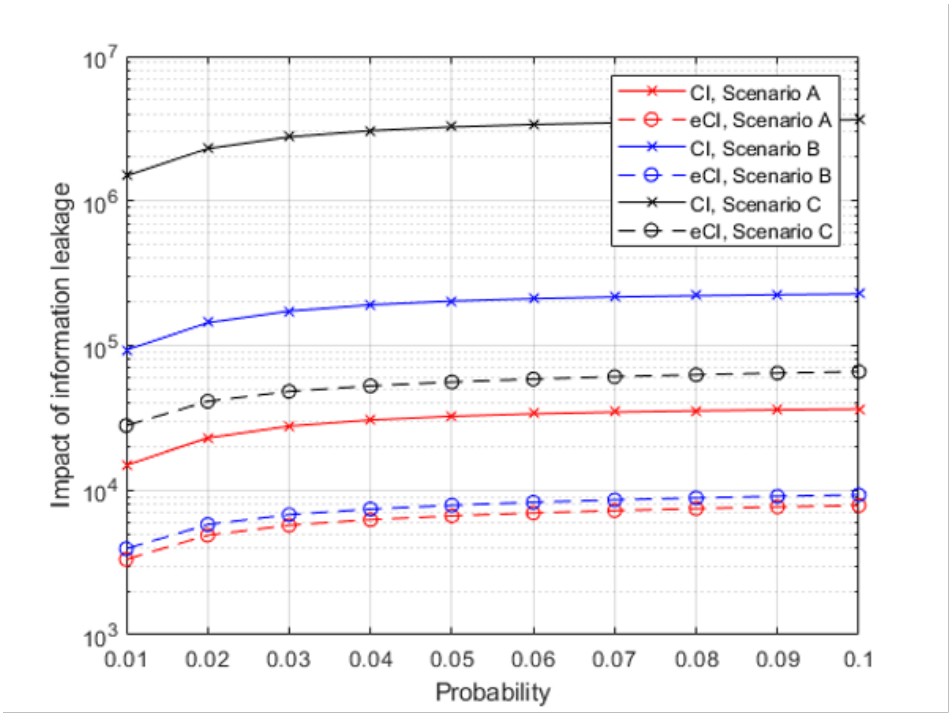

FIGURE 4.12 Scenario-based evaluation of impact of information leakage in eCI.

Fig. 4.12 shows the performance evaluation results of the CI and eCI systems based on the scenario-based simulation. The results demonstrate

that in all scenarios, the eCI system significantly reduces the impact of information leakage compared to the conventional CI system. Moreover, as the total number of CIs used per user (i.e., $\frac{u}{s}$) increases, the resulting impact of CI leakage also increases. Furthermore, the impact of leakage becomes more pronounced as the attacker compromises a greater portion of service s, which varies depending on the security level of each scenario.

3) Complexity Analysis and Security-Complexity Adaptation

TABLE 4.7
Comparison of space complexity.

	CI system	eCI system
Space complexity	$O(n)$	$O(n^2)$

① Space complexity

In an eCI system, as multiple CIs of a data subject are generated depending on the service, the space complexity for CI storage and management increases. Table 4.7 presents a comparison of the space complexity of CI and eCI systems. The conventional CI system has a space complexity of $O(n)$ because it stores only one CI per data subject. Moreover, the eCI system has a space complexity of $O(n^2)$ because it stores the same number of CIs per data subject as there are services. Although the eCI system requires more memory space for CI management than the conventional method, dynamic CI generation and updating reduces the impact of CI leakage and improves the security of services using digital identity.

② Communication complexity

The eCI system generates multiple CIs and updates them periodically, which increases the communication complexity and degrades the data transmission performance when compared with the conventional CI system, in which a fixed and single CI is maintained for each user. When

all users use eCI, the communication complexity and management costs significantly increase. This can be adjusted by applying the security level of the CIs differently, considering the importance of the user. If a user is divided into two levels, namely, an administrator and a general user, the general user can adaptively manage their complexity by reducing the number of CI generations and updates, whereas the CIs of the administrator are strictly managed by the eCI system. Table 4.8 presents the parameters and values of the complexity evaluation simulation of the adaptive eCI.

TABLE 4.8
Parameters for the complexity evaluation of adaptive CI.

Parameter	eCI	CI	
		Administrator	General user
# of total transaction	10,000	10,000	10,000
# of total users	10,000	10,000	10,000
# of CIs per user	[10:5:50]	50	5
CI update interval	[50, 100, 200]	[50, 100, 200]	200

Communication complexity (C_{comm}) is the ratio of transactions generated by CI updates to the total transactions during the entire service period and has a value between zero and one (Eq. 4.8).

$$C_{comm} = 1 - \frac{K - \frac{K}{j} \times s}{K}, \tag{4.8}$$

Fig. 4.13 presents the communication complexity of the eCI system, wherein the number of CIs for all users and the CI update interval are

the same, irrespective of their importance. Fig. 4.13 presents an analysis of the communication complexity when the number of CIs per user increases from 10 to 50, with the CI update interval set to 50, 100, and 200. As can be seen from the figure, the communication complexity increased as the number of CIs and update intervals increased.

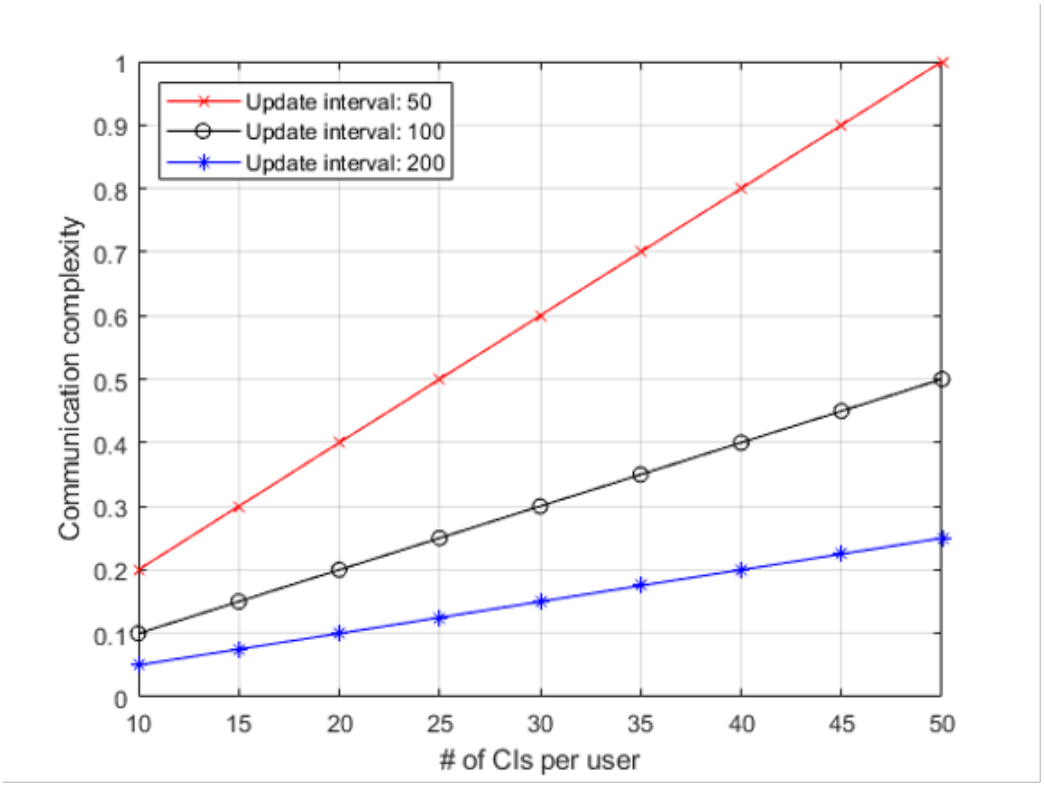

FIGURE 4.13 Communication complexity of enhanced connecting information systems with respect to the numbers of connecting information per user.

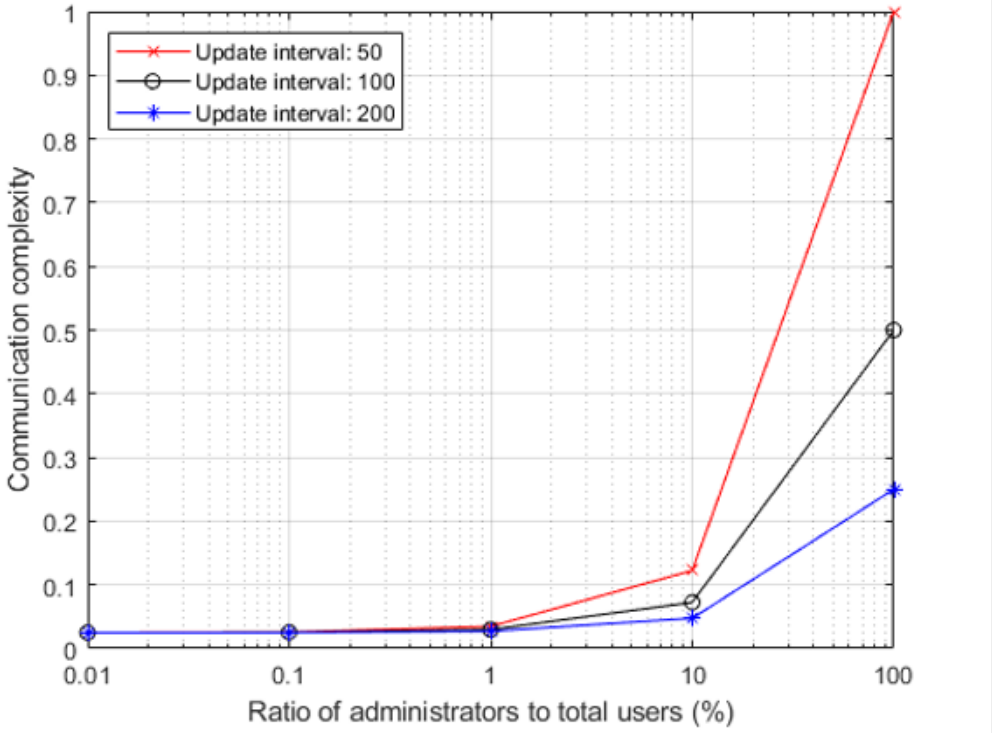

FIGURE 4.14 Communication complexity of adaptive enhanced connecting information systems with respect to the ratio of administrators to total users.

Fig. 4.14 illustrates the communication complexity when the ratio of administrators to total users is increased in an adaptive eCI system, wherein the number of CIs and update intervals are set differentially according to the importance of the user role. The number of CIs for administrators and general users was fixed at 50 and 5, respectively, and the CI update interval was increased to 50, 100, and 200 for administrators and 200 for general users. In the adaptive eCI system, when 10% and 90% of the total users were administrators and general

users, respectively, the maximum communication complexity was 0.1225 (a reduction of 87.75% when compared with the eCI system, where the communication complexity was measured as 1), although both systems exhibited the same parameter values (i.e., 50 CIs per user and an update interval of 50).

An eCI system can enhance security by reducing the amount of information leakage by generating several digital IDs for each user and periodically updating them accordingly. However, this method is subject to an increased complexity. The eCI system can achieve lower complexity while maintaining security by adaptively adjusting the security level according to the importance of the user role or authority instead of assigning an equally high security level to all users.

TABLE 4.9
Comparison of CI, DID and eCI systems in terms of performance, security, and management.

Category	CI	DID	eCI
Performance	Minimal overhead and latency due to static ID usage.	Latency may occur depending on the performance of underlying distributed storage technologies, such as blockchain.	Potential performance degradation due to ID update latency and synchronization errors.
Security	Vulnerable to privacy breaches and traceability	Similar to CI, DID also relies on unique identifiers,	Periodic ID updates significantly

	due to the use of fixed identifiers.	making it equally vulnerable to identity leakage and tracking.	reduce the risk of personal information leakage and traceability in the event of ID compromise.
Management	Easiest to manage, as all IDs are centrally stored and reused for identification and authentication.	Similar to the CI system in that it uses static identifiers but requires an additional converter to replace CI with DID, resulting in extra operational overhead and potential security risks during the conversion process.	Requires additional management overhead due to continuous CI updates. However, update frequency and the number of CIs can be adjusted according to security levels, allowing for cost optimization. Also maintains compatibility with existing CI structures, offering advantages in interoperability.

Table 4.9 provides a comparison of the eCI system, the conventional CI system, and a representative identity management approach—DID—in terms of performance, security, and management overhead. Compared to

the eCI system, which involves delays due to periodic ID updates, both the CI and DID systems incur minimal overhead and latency. For DIDs, latency varies depending on the underlying distributed ledger or blockchain technology. However, both CI and DID use unique, static identifiers, making them vulnerable to privacy breaches and traceability risks in the event of identity leakage. In terms of management, the CI system requires the least overhead, as it stores and reuses a fixed identifier for both identification and authentication. Similarly, DID also uses a static identifier; however, to replace existing CI-based digital identities, it requires the deployment of an additional converter to translate CI into DID format, which introduces extra operational complexity and cost. Moreover, the conversion process itself may introduce additional security risks, such as exposure during format translation or vulnerabilities in the converter module, particularly if the process is not adequately protected. The eCI system, on the other hand, incurs higher management overhead due to the need for continuous CI updates. Nevertheless, it allows for flexibility in adjusting the update interval and number of CIs based on the desired security levels, enabling cost optimization. Furthermore, because the eCI system retains compatibility with the existing CI structure, it offers advantages in terms of integration and interoperability. In summary, the eCI system reduces the risks of personal information leakage and traceability associated with the use of fixed identifiers in conventional CI systems while also offering better compatibility with existing CI infrastructure compared to approaches such as DID.

5. Discussion and Future Work

1) Discussion

Given that the eCI systems periodically update the CI, ensuring backward compatibility with legacy CI systems is of critical importance. To maintain interoperability between the eCI and conventional CI systems, a CI type indicator can be exchanged during the user identification process to enable digital identity mode detection. For example, during the issuance of a MyData access token, the system may first negotiate the CI type and then proceed with user identification using either the legacy CI or the eCI, depending on system capabilities. Service providers equipped to support the eCI framework can offer enhanced security features, whereas those without such support can ensure interoperability by detecting the identity mode and reverting to the conventional CI-based procedures.

In the event of synchronization failure during the CI update process, relying parties only need to retain the seed value and the time interval parameter to recompute the time-based counter and the corresponding *PRN*, thereby reconstructing a valid eCI. The eCI can be deterministically generated based on the current timestamp, the predefined interval, and seed value, enabling on-demand synchronization without requiring continuous updates. Furthermore, if the eCI is compromised, the CI-issuing authority can provide a new secret key and seed value, allowing secure revocation and replacement of the compromised eCI, thereby mitigating the risk of information leakage.

2) Future work

The eCI system addresses the vulnerability of static identifier exposure by introducing a mechanism that periodically updates digital identifiers derived from the conventional CI. Notably, the eCI system maintains compatibility with existing CI infrastructures by reusing the same cryptographic structure, which gives it a significant advantage over DID-based digital identity solutions that require the deployment of additional CI-to-DID conversion systems.

The cryptographic design of the eCI system combines the original CI generation scheme with an OTP-based update mechanism. The security of the eCI system depends on minimizing the vulnerabilities in both the CI generation and update processes. For instance, if the secret key used during CI generation or the seed value used in the update process is compromised, attackers may be able to recreate the same CI, thereby undermining the system's integrity. To prevent such threats, it is essential to use secure communication protocols such as TLS 1.3 during the CI generation process and implement the OTP system securely through hardware tokens, encrypted sessions, and secure communication channels.

In addition, since the identifier is updated at regular intervals, there is a risk of tracking attacks where an adversary monitors updated identifiers over time to infer user behavior or activity patterns. To address this, we aim to construct threat scenarios involving identifier tracking and develop a privacy leakage analysis framework to evaluate the potential impact of such attacks on user privacy.

The eCI system introduces a foundational concept of updating fixed digital identifiers and can be applied not only to CI management but also to various digital identity use cases across industries. Specifically, the RN value in the CI can be replaced with a range of digital identity values to construct a general-purpose digital ID update mechanism. This framework can be applied to application-layer services such as electronic voting systems and healthcare platforms, as well as e-government services like mobile ID cards and digital driver's licenses. It can also be utilized for securing identity in communication systems and IoT devices.

Furthermore, the eCI system can be integrated with hardware-based identification and authentication mechanisms. For instance, it can leverage physical unclonable functions (PUFs) as hardware fingerprints for device identification and authentication while enabling periodic ID updates at the application level. A hardware-based implementation of the eCI system can improve security by eliminating the need for pre-shared seed values, relying instead on device-specific hardware fingerprints. In addition, by offloading computational tasks to hardware, this approach can reduce processing overhead, minimize latency, and effectively address the complexity challenges associated with software-based identity management. As part of our future work, we plan to design and implement a hardware-integrated eCI authentication and identification framework.

6. Conclusion

CI is an emerging digital identity technology that is widely used in various industries to verify the identity of data subjects and individuals without directly using unique identification numbers such as the resident registration number. However, for system convenience, all CI-using organizations store and use the same CI for the data subject. Consequently, CI leakage leads to the leakage of personal and sensitive information. This study proposes a technology that generates unique identification numbers of data subjects into multiple digital identities according to the service used by the data subject and updates the digital identities at regular intervals. The proposed eCI system generates and periodically updates the CIs of data subjects depending on the organization and time. Moreover, the CI is updated at regular intervals, which minimizes the scale of the damage caused by personal information leakage and its impact. Furthermore, the security evaluation of the proposed eCI system confirmed improved security by reducing the impact of CI leakage when compared with the conventional CI system. An illustrative method for an adaptive eCI system was proposed to improve complexity, which adaptively adjusts the security level according to the importance or authority of the data subject. The eCI system methodology proposed in this study can strengthen the security of existing centralized digital identity systems without significantly increasing their complexity. The eCI system can therefore serve as a key technology supporting the

core systems of e-government, thereby ensuring both privacy and data sovereignty. Future research should focus on investigating methods to improve the complexity and stability of eCI systems.

Table of Notations (Chapter IV)

Notations	Description
RN	Resident registration number
S_A, sk	Secret keys
$Temp$	Encrypted RN
SI	Service identifier
C_T	Counter value
T_0	Unix time
T_C	Current time
T_I	Time interval
d	Digit
I	Impact of information leakage
K	Service period
t	Transaction
u	The number of data subjects
p	The probability of CI leakage in one transaction
i	CI update cycle
j	The number of CI updates
w_t	CI leakage time weight
r_t	CI round at time point t
q	The proportion of compromised services
α	The synchronization failure rate

Ⅴ. Conclusion

This dissertation proposed an integrated security framework that combines physical and logical techniques to enhance the security of financial transactions and user authentication. Specifically, a PLKG scheme based on wireless communication was introduced to address the limitations of traditional short-range payment methods such as NFC. By leveraging the inherent randomness of wireless channels, the proposed PLKG method enables the generation of cryptographic keys without relying on preshared secrets or centralized trust. This approach not only extends communication coverage and improves data transmission speed via Wi-Fi-based communication but also enhances security by protecting MAC headers and supporting low-latency encryption of sensitive payment data. Moreover, it demonstrates robustness and reliability in untrusted relay environments, thereby offering high security and broad applicability in real-world financial infrastructures.

In addition, this study proposes a dynamic digital identity protection system as the logical security layer. Unlike static identifiers, which are vulnerable to leakage and reuse attacks, the proposed system periodically generates and updates digital identities. It reduces the risk of information leakage by enabling time-variant identity values and maintains backward compatibility with legacy CI systems through mode detection mechanisms. A complexity management strategy is also incorporated to address the overhead introduced by frequent ID updates and multi-ID handling, ensuring practical applicability in large-scale service

environments.

By combining these two complementary approaches—physical-layer key generation and dynamic digital identity protection—this work presents a comprehensive security architecture capable of addressing the evolving threats associated with financial services in the era of digital transformation. The proposed system enhances communication security, safeguards user identity, and supports the development of a trustworthy and resilient digital economy.

ACKNOWLEDGEMENTS

The author would like to express her sincere gratitude to Professor Il-Gu Lee for his invaluable guidance and supervision throughout this study, as well as to Professors Heejung Yu and Jung Hoon Lee for their co-supervision, insightful advice, and continuous support.

References

Abdalhafid, A.A., Subramaniam, S.K., Zukarnain, Z.A., Ayob, F.H., 2024. Multi-link operation in IEEE802.11be extremely high throughput: a survey. IEEE Access. 12, 46891- 46906. https://doi.org/10.1109/ACCESS.2024.3378997.

Abdelazeem, I. et al., 2024. A lossless quantization approach for physical-layer key generation in vehicular ad hoc networks based on received signal strength. Veh. Commun., vol. 49, p. 100809. https://doi.org/10.1016/j.vehcom.2024.100809.

Abyaneh, A.Z. et al., 2023. Empowering next-generation IoT WLANs through blockchain and 802.11ax technologies. IEEE Trans. Intell. Transp. Syst., vol. 24, no. 7, 7412-7421. https://doi.org/10.1109/TITS.2022.3153484.

Ahmed, N. et al., 2022. MAC protocols for IEEE 802.11ah-based Internet of Things: A survey. IEEE Internet Things J., vol. 9, no. 2, pp. 916-938. https://doi.org/10.1109/JIOT.2021.3104388.

Alam, M. et al., 2024. Analyzing the suitability of IEEE 802.11ah for next generation Internet of Things: A comparative study. Ad Hoc Netw., vol. 156, p. 103437. https://doi.org/10.1016/j.adhoc.2024.103437.

Aldaghri, N. and Mahdavifar, H., 2020. Physical layer secret key generation in static environments. IEEE Trans. Inf. Forensics Sec., vol. 15, pp. 2692-2705. https://doi.org/10.1109/TIFS.2020.2974621.

Alrawad, M., Lutfi, A., Almaiah, M.A., Elshaer, I.A., 2023. Examining the influence of trust and perceived risk on customers intention to use NFC

mobile payment system. J Open Innov Technol Mark Complex. 9, 10007 0. https://doi.org/10.1016/j.joitmc.2023.100070.

Alsabah, M., Naser, M.A., Mahmmod, B.M., Abdulhussain, S.H., Eissa, M. R., Al-Baidhani, A., et al.,2021. 6G wireless communications networks: a comprehensive survey. IEEE Access. 9, 148191-148243. https://doi.org/1 0.1109/ACCESS.2021.3124812.

Alwazzeh M, Karaman S, Shamma MN. Man in the middle attacks against SSL/TLS: mitigation and defeat. JCSANDM 2020;9:449-68. https://doi.or g/10.13052/jcsm2245-1439.933

Anthes G. Estonia: a model for e-government. Commun ACM 2015;58:18-2 0. https://doi.org/10.1145/2754951.

Assaf, T., Al-Dweik, A., Iraqi, Y., Jangsher, S., Pandey, A., Giacalone, J.P., et al.,2023. High-rate secret key generation using physical layer securit y and physical unclonable functions. IEEE Open J Commun Soc. 4, 209 -225. https://doi.org/10.1109/OJCOMS.2023.3234338.

Aubert B, Chan Y. Evolving strategic IS themes. J Strateg Inf Syst 2024;3 3:101821. https://doi.org/10.1016/j.jsis.2024.101821

Bossenko I, Piho G, Ross P. Modelling a Patient Identifier System in the E stonian National Health Information System. In: International Conference on Digital Economy. Cham: Springer Nature Switzerland;2024. pp.132-14 5. https://doi.org/10.1007/978-3-031-76368-7_10

Dai, D., An, Z., Pan, Q., Yang ,L., 2024. Harnessing NFC to generate stand ard optical barcodes for NFC-missing smartphones. IEEE Trans Mob C omput. 23, 12952-12968. https://doi.org/10.1109/TMC.2024.3420410.

Deng, C. et al., 2020. IEEE 802.11be wi-fi 7: New challenges and opportuni

ties. IEEE Commun. Surv. Tutor., vol. 22, no. 4, pp. 2136-2166. https://doi.org/10.1109/COMST.2020.3012715.

Fanari, L. et al., 2024. Optimized IEEE 802.11ax for smart warehouses. Ad Hoc Netw., vol. 158, p. 103415. https://doi.org/10.1016/j.adhoc.2024.103415.

Furqan, H.M., Hamamreh, J.M., Arslan, H., 2021. New physical layer key generation dimensions: subcarrier indices/positions-based key generation. IEEE Commun Lett. 25, 59-63. https://doi.org/10.1109/LCOMM.2020.3025262.

Galati-Giordano, L., Geraci, G., Carrascosa, M., Bellalta, B., 2024. What will Wi-Fi 8 Be? A Primer on IEEE 802.11bn Ultra High Reliability. IEEE Commun Mag. 62, 126-132. https://doi.org/10.1109/MCOM.001.2300728.

Gan W, Ye Z, Wan S, Yu PS. Web 3.0: The future of internet. In: Companion Proceedings of the ACM Web Conference 2023. Austin, TX, USA;2023:1266-1275. https://doi.org/10.1145/3543873.3587583

Gao, N., Han, Y., Li, N., Jin, S., Matthaiou, M., 2024. When physical layer key generation meets RIS: opportunities, challenges, and road ahead. IEEE Wirel Commun. 31, 355-361. https://doi.org/10.1109/MWC.013.2200538.

Garcia-Rodriguez, A., López-Pérez, D., Galati-Giordano, L., Geraci, G., 2021. IEEE 802.11be: wi-fi 7 strikes back. IEEE Commun Mag. 59, 102-108. https://doi.org/10.1109/MCOM.001.2000711.

Guo, D., Cao, K., Xiong, J., Ma, D., Zhao, H., 2021. A lightweight key generation scheme for the internet of things. IEEE Internet Things J. 8, 12137-12149. https://doi.org/10.1109/JIOT.2021.3060438.

Han, B. et al., 2023. FLoRa: Sequential fuzzy extractor based physical layer key generation for LPWAN. Future Gener. Comput. Syst., vol. 140, pp. 253-265. https://doi.org/10.1016/j.future.2022.10.018.

Jeon SE, Lee YJ, Lee IG. Software Defined Range-Proof Authentication Me chanism for Untraceable Digital ID. CMES 2025;142;3213-28. https://doi. org/10.32604/cmes.2025.062082

Jeon, Y.R. et al., 2024. ART: Adaptive relay transmission for highly reliable communications in next-generation wireless LANs," Comput. Netw., vol. 254, p. 110752. https://doi.org/10.1016/j.comnet.2024.110752.

Kang J, Jang YY, Kim J, Han SH, Lee KR, Kim M, Eom JS. South Korea 's responses to stop the COVID-19 pandemic. American Journal of Infe ction Control 2020;48:1080-1086. https://doi.org/10.1016/j.ajic.2020.06.003

Kazaz, T., Janssen, G.J.M., Romme, J., van der Veen, A.J,. 2022. Delay esti mation for ranging and localization using multiband channel state infor mation. IEEE Trans Wirel Commun. 21, 2591-25607. https://doi.org/10.1 109/TWC.2021.3113771.

Keshavarzi, M. et al., 2024. A new practical physical layer secret key gene ration in the presence of an untrusted relay. Phys. Commun., vol. 66, p. 102407. https://doi.org/10.1016/j.phycom.2024.102407.

Khan HU, Sohail M, Nazir S, Hussain T, Shah B, Ali F. Role of authentic ation factors in Fin-tech mobile transaction security. Journal of Big Dat a 2023;10:138. https://doi.org/10.1186/s40537-023-00807-3

Khanh, Q.V., Hoai, N.V., Manh, L.D., Le, A.N., Jeon, G., 2022. Wireless co mmunication technologies for IoT in 5G: vision, applications, and challen ges. Wirel Commun Mob Comput. 2022, 3229294. https://doi.org/10.1155/

2022/3229294.

Kim H, Park KW, Choi D, Lee Y. Estimating resident registration numbers of individuals in Korea: revisited. KSII Trans Internet Inf Syst 2018;12: 2946-59. https://doi.org/10.3837/tiis.2018.06.027

Kim HJ, Kim S, Lee S. On privacy enhancement using u-distinguishability to COVID-19 contact tracing approach in Korea. In: Kose U, Gupta D, de Albuquerque VHC, Khanna A, editors. Data science for COVID-19. Elsevier; 2021. p. 661-73. https://doi.org/10.1016/B978-0-12-824536-1.00010-1

Kong, Y. et al., 2018. The security network coding system with physical la yer key generation in two-way relay networks," IEEE Access, vol. 6, p p. 40673-40681. https://doi.org/10.1109/ACCESS.2018.2858282.

Kulkarni, R.D., 2020. Near field communication (NFC) technology and its a pplication, in: Pawar, P.M., Balasubramaniam, R., Ronge, B.P., Salunkhe, S.B., Vibhute, A.S., Melinamath, B., editors Techno-Soc., Cham: Springe r International Publishing; 2021, pp. 745-751. https://doi.org/10.1007/978-3-030-69921-5_74.

Lee D, Choi B. Policies and innovations to battle Covid-19 - a case study of South Korea. Health Policy Technol 2020;9:587-97. https://doi.org/10.1016/j.hlpt.2020.08.010

Lee J, Lee H, Jeong J, Kim D, Kwon TT. Analyzing spatial differences in the TLS security of delegated web services. Acad Med 2021:475-87. https://doi.org/10.1145/3433210.3453107

Lee S, Shin Y, Hur J. Return of version downgrade attack in the era of TLS 1.3. CoNEXT '20: the, 1(3) 16th International Conference on Emergin

g Networking Experiments and Technologies Nov 24; 2020. p. 157-68. https://doi.org/10.1145/3386367.3431310.

Lee SJ, Lee JM, Lee IG. Low latency and secure data encryption for multi-hop biometric authentication in distributed networks. Internet of Things 2025;30;101501. https://doi.org/10.1016/j.iot.2025.101501

Lee, W., Baek, S.Y., Kim, S.H., 2021. Deep-learning-aided RF fingerprinting for NFC security. IEEE Commun Mag. 59, 96-101. https://doi.org/10.1109/MCOM.001.2000912.

Lin, R., Xu, L., Fang, H., Huang, C., 2020. Efficient physical layer key generation technique in wireless communications. EURASIP J Wirel Commun Netw. 2020,13. https://doi.org/10.1186/s13638-019-1634-7.

Liu, C. et al., 2023. IRS-aided secure communications over an untrusted AF relay system. IEEE Trans. Wirel. Commun., vol. 22, no. 12, pp. 8620-8633. https://doi.org/10.1109/TWC.2023.3264626.

Liu Y, He D, Obaidat MS, Kumar N, Khan MK, Raymond Choo KK. Blockchain-based identity management systems: a review. J Netw Comput Appl 2020;166:102731. https://doi.org/10.1016/j.jnca.2020.102731.

Lopez-Perez, D. et al., 2019. IEEE 802.11be extremely high throughput: The next generation of wi-fi technology beyond 802.11 ax. IEEE Commun. Mag., vol. 57, no. 9, pp. 113-119. https://doi.org/10.1109/MCOM.001.1900338.

López-Raventós, A. and Bellalta, B., 2022. Multi-link operation in IEEE 802.11be WLANs. IEEE Wirel. Commun., vol. 29, no. 4, pp. 94-100. https://doi.org/10.1109/MWC.006.2100404.

Lu, T., Chen, L., Zhang, J., Cao, K., Hu, A., 2022. Reconfigurable intelligent

surface assisted secret key generation in quasi-static environments. IEE E Commun Lett. 26, 244-248. https://doi.org/10.1109/LCOMM.2021.31306 35.

Luo, H. et al., 2023. Joint secure transceiver design for an untrusted MIMO relay assisted over-the-air computation networks with perfect and impe rfect CSI. IEEE Trans. Inf. Forensics Sec., vol. 18, pp. 2508-2523. https: //doi.org/10.1109/TIFS.2023.3266928.

Manoj T, Makkithaya K, VG N. A Blockchain Based Decentralized Identifie rs for Entity Authentication in Electronic Health Records. Cogent Engin eering 2022;9. https://doi.org/10.1080/23311916.2022.2035134

Meng W, Zhu L, Li W, Han J, Li Y. Enhancing the security of FinTech a pplications with map-based graphical password authentication. Future G eneration Computer Systems 2019;101:1018-1027. https://doi.org/10.1016/j. future.2019.07.038

Moradi H, Vaezi, A. Lessons learned from Korea: COVID-19 pandemic. Inf ection Control & Hospital Epidemiology 2020;41:873-874. https://doi.org/1 0.1017/ice.2020.104

Norrman K, Näslund M, Dubrova E. Protecting IMSI and user privacy in 5 G networks. In: Proceedings of the 9th EAI international conference on mobile multimedia communications. 2016. pp. 159-166.

OWASP. Broken object level authorization. https://owasp.org/API-Security/ editions/2023/en/0xa1-broken-object-level-authorization/, 2023a (accessed 28 July 2024)

OWASP. Cross site request forgery (CSRF). https://owasp.org/www-comm unity/attacks/csrf. 2023b (accessed 28 July 2024)

Palamà I, Gringoli F, Bianchi G, Blefari-Melazzi N. (2021). IMSI Catchers i n the wild: A real world 4G/5G assessment. Comput Netw 194;108137. https://doi.org/10.1016/j.comnet.2021.108137

Perugini L, Vesco A. On the integration of Self-Sovereign Identity with T LS 1.3 handshake to build trust in IoT systems. Internet of Things 202 4;25;101103. https://doi.org/10.1016/j.iot.2024.101103

Ramamoorthi K, Stamenova V, Liu RH, Bhattacharyya O. The Implementat ion of Federated Digital Identifiers in Health Care: Rapid Review. Journ al of Medical Internet Research 2024;26:e45751. https://doi.org/doi:10.2196 /45751

Saeed RA, Saeed MM, Mokhtar RA, Alhumyani H, Abdel-Khalek S. Pseud onym Mutable Based Privacy for 5G User Identity. Comput Syst Sci E ng 2021;39. https://doi.org/10.32604/csse.2021.015593

Saha, R.K., 2021. Coexistence of cellular and IEEE 802.11 technologies in u nlicensed spectrum bands-a survey. IEEE Open J. Commun. Soc., vol. 2, pp. 1996-2028. https://doi.org/10.1109/OJCOMS.2021.3106502.

Sambra AV, Mansour E, Hawke S, Zereba M, Greco N, Ghanem A, et al. Solid: A platform for decentralized social applications based on linked d ata, MIT CSAIL, Qatar computing Research Institute; 2016. https://api.s emanticscholar.org/CorpusID:49564404

Samir E, Wu H, Azab M, Xin C, Zhang Q. DT-SSIM: A decentralized trus tworthy self-sovereign identity management framework. IEEE Internet o f Things Journal 2021;9:7972-7988. https://doi.org/10.1109/JIOT.2021.3112 537

Satybaldy A, Ferdous MS, Nowostawski M. A taxonomy of challenges for

self-sovereign identity systems. IEEE Access 2024;12;16151-16177.https://doi.org/10.1109/ACCESS.2024.3357940

Senyo PK, Karanasios S, Komla Agbloyor EK, Choudrie J. Government-Led digital transformation in FinTech ecosystems. J Strateg Inf Syst 2024;33:101849. https://doi.org/10.1016/j.jsis.2024.101849

Shahiri, V., Behroozi, H., Kuhestani, A., Wong, K.K., 2024. Reconfigurable-intelligent-surface-assisted secret key generation under spatially correlated channels in quasi-static environments. IEEE Internet Things J. 11,15808-15822. https://doi.org/10.1109/JIOT.2023.3349354.

Stockburger L, Kokosioulis G, Mukkamala A, Mukkamala RR, Avital M. Blockchain-enabled decentralized identity management: The case of self-sovereign identity in public transportation. Blockchain: Research and Applications 2021;2:100014. https://doi.org/10.1016/j.bcra.2021.100014

Sun, R. et al., 2023. Covertness and secrecy study in untrusted relay-assisted D2D networks. IEEE Internet Things J., vol. 10, no. 1, pp. 17-30. https://doi.org/10.1109/JIOT.2022.3201021.

Szott, S. et al., 2022. Wi-fi meets ML: A survey on improving IEEE 802.11 performance with machine learning. IEEE Commun. Surv. Tutor., vol. 24, no. 3, pp. 1843-1893. https://doi.org/10.1109/COMST.2022.3179242.

Thai, C.D.T. et al., 2016. Physical-layer secret key generation with colluding untrusted relays. IEEE Trans. Wirel. Commun., vol. 15, no. 2, pp. 1517-1530. https://doi.org/10.1109/TWC.2015.2491935.

Tian, L. et al., 2021. Wi-fi HaLow for the Internet of Things: An up-to-date survey on IEEE 802.11ah research," J. Netw. Comput. Appl., vol. 182, May, p. 103036. https://doi.org/10.1016/j.jnca.2021.103036.

Usman, M., Althunibat, S., Qaraqe, M., 2022. A channel state information-based key generation scheme for internet of things. Secur Commun Netw. 2022, 7976319. https://doi.org/10.1155/2022/7976319.

Waleed M, Skouby KE, Kosta S. Decentralized Identifiers for IoT Systems: Current Status and Future Vision. In: International Symposium on Distributed Computing and Artificial Intelligence. Cham: Springer Nature Switzerland;2023. pp. 408-417. https://doi.org/10.1007/978-3-031-38318-2_40

Xiao, Q. et al., 2025. Optimal subcarrier allocation scheme for physical-layer key generation in an OFDMA network. IEEE Trans. Inf. Forensics Sec., pp. 1-1. https://doi.org/10.1109/TIFS.2025.3566242.

Xu, H., Zhang, J., Tang, P., Tian, L., Wang, Q., Liu, G., 2024. An empirical study on channel reciprocity in TDD and FDD systems. IEEE Open J Veh Technol. 5, 108-124. https://doi.org/10.1109/OJVT.2023.3339799.

Xu, P. et al., 2024. Physical-layer secret and private key generation in wireless relay networks with correlated eavesdropping channels. IEEE Trans. Inf. Forensics Sec., vol. 19, pp. 985-1000. https://doi.org/10.1109/TIFS.2023.3329740.

Yang, M. et al., 2021. MAC Technology of IEEE 802.11ax: Progress and Tutorial. Mob. Netw. Appl., vol. 26, no. 3, pp. 1122-1136, 2021, https://doi.org/10.1007/s11036-020-01622-3.

Yang X, Li W. A zero-knowledge-proof-based digital identity management scheme in blockchain. Comput Sec 2020;99:102050. https://doi.org/10.1016/j.cose.2020.102050

Yi, P. et al., 2025. RIS-assisted seamless connectivity in wireless multi-hop relay networks. IEEE Trans. Mob. Comput., pp. 1-13. https://doi.org/10.

1109/TMC.2025.3557676.

Yin J, Xiao Y, Pei Q, Ju Y, Liu L, Xiao M, Wu C. SmartDID: a novel privacy-preserving identity based on blockchain for IoT. IEEE Internet of Things Journal 2022;10:6718-6732. https://doi.org/10.1109/JIOT.2022.3145089

Yu, H., Kim, T., 2019. Training and data structures for AN-aided secure communication. IEEE Syst J. 13, 2869-2872. https://doi.org/10.1109/JSYST.2018.2859446.

Yu, H., Lee, H., 2020. Joint optimization of power and fronthaul compression for data and pilot signals in uplink C-RANs. IEEE Syst J. 14, 4990-5001. https://doi.org/10.1109/JSYST.2020.2980164.

Yu Y, Au MH, Ateniese G, Huang X, Susilo W, Dai Y, Min G. Identity-based remote data integrity checking with perfect data privacy preserving for cloud storage. IEEE Trans. Inform. Forensic Secur 2017;12:767-78. https://doi.org/10.1109/TIFS.2016.2615853

Zhang, B., Waqas, M., Tu, S., Hussain, S., Rehman, S., 2021. Power allocation strategy for secret key generation method in wireless communications. Comput Mater Contin. 68, 2179-2188. https://doi.org/10.32604/cmc.2021.016553.

Zhang, J., Rajendran, S., Sun, Z., Woods, R., Hanzo, L., 2019. Physical layer security for the internet of things: authentication and key generation. IEEE Wirel Commun. 26, 92-98. https://doi.org/10.1109/MWC.2019.1800455.

Zhang, S., Zhu, D., Liu, Y., 2024. Artificial intelligence empowered physical layer security for 6G: State-of-the-art, challenges, and opportunities. Comput Netw. 242, 110255. https://doi.org/10.1016/j.comnet.2024.110255.

논문개요

금융 보안을 위한 무선 통신 및 디지털 신원 보안 기술

박소현
미래융합기술공학과
성신여자대학교 대학원

정보통신 기술의 급속한 발전에 따라 금융, 의료, 정부 시스템 등 다양한 분야에서 디지털 전환이 활발히 진행되고 있다. 금융 분야에서는 실물 카드를 기반으로 한 전통적인 결제 방식에서, NFC (near field communication)과 같은 근거리 통신 기술을 활용한 비접촉 결제 시스템으로 발전하였다. 또한, 자기주권 신원(self-sovereign identity, SSI) 기술의 발전으로 모바일 신원 시스템이 실물 신분증을 대체하는 수단으로 채택되고 있으며, 마이데이터와 같은 데이터 기반 금융 서비스는 사용자 중심의 개인정보 통제를 가능하게 하고 있다.

하지만, 금융 서비스의 디지털화에 따라 무선 통신 및 사용자 인증 과정에서의 보안 위협이 확대되고 있다. 본 논문에서는 금융 서비스의 거리적 한계점을 극복하고 통신 과정에서 발생할 수 있는 취약점을 해결하기 위한 물리계층 비밀키 생성(PLKG) 기반의 무선랜 금융 결제 시스템을 제안한다. 더불어, 식별 및 인증 과정 중 발생할 수 있는 개인정보 유출 위협을 완화하기 위한 디지털 신원 보호 메커니즘을 제안한다. 본 학위 논문에서 제안하는 물리 계층과 논리 계층의 보안 기법을 통합하는 시스템을 통하여, 디지털 금융 서비스를 위한 통신 및 사용자 관리 시스템의 안정성과 신뢰성을 향상시킬 수 있다.