SecCSI: Securing Wireless Environment With RIS Against CSI-Forgery Attacks

Yicheng Liu, *Graduate Student Member, IEEE*, Zhao Li, *Senior Member, IEEE*, Kang G. Shin, *Life Fellow, IEEE*, Zheng Yan, *Fellow, IEEE*, Jia Liu, *Senior Member, IEEE*, and Siwei Le

***Abstract*—Channel state information (CSI) is known to be crucial for both enhancing the transmission performance and ensuring physical-layer security (PLS) in wireless communication systems. To estimate a channel's CSI, the transmitter (Tx) typically broadcasts a predetermined pilot signal, then the receiver (Rx) computes the channel coefficients based on the received pilot signal and returns the estimated CSI to the Tx. Most, if not all, of existing communication algorithms simply assume that the fed-back CSI is reliable/secure. However, in practice, a malicious terminal may send falsified CSI to the infrastructure, thus compromising the throughput and/or security of the communication over the channel. Although some researchers have already identified this vulnerability, demonstrated the feasibility of the CSI-forgery attacks, and designed countermeasures thereof, their methods either i) are tailored to specific types of attacks, thus lacking generality, or ii) require modifications to the pilot sequence and hence the protocol. To counter the CSI-forgery attacks and remove/mitigate the deficiencies of existing countermeasures, we first develop a comprehensive CSI-forgery model that can subsume the existing CSI-forgery attacks as special instances to facilitate the design of general countermeasures. Then, we propose a novel approach, called `SecCSI`, to detect potential CSI-forgery activities and identify their initiators using reconfigurable intelligent surface (RIS). `SecCSI` leverages the RIS to secretly and dynamically modify the wireless environment transparently to the receiver (Rx) in which the pilot signal is transmitted. The infrastructure can, therefore, detect any attempted manipulation of CSI by appropriately configuring the reflection coefficient matrix of the RIS, transmitting the pilot signal, and analyzing all CSI feedback. `SecCSI` can serve as a guard module for existing communication systems that simply accept the fed-back CSI without checking its trustworthiness. Our theoretical analysis, experimental and numerical evaluations have shown `SecCSI` to effectively detect the CSI-forgery attacks and identify the attacker.**

***Index Terms*—Channel state information (CSI), physical-layer security (PLS), CSI-forgery attacks, reconfigurable intelligent surface (RIS).**

I. Introduction

With the rapid advancement of wireless communication technologies, wireless networks are evolving towards ultra-high capacity, ultra-densification, and high-bandwidth/speed services. Wireless signal propagation is known to be sensitive to channel condition. To counter or compensate for channel fading, acquiring the channel state information (CSI), therefore, becomes crucial. This enables the design of efficient transmission mechanisms that can match the transmitted signals with the characteristics of a wireless channel and manage the interference among these signal transmissions, thereby achieving a high data rate and avoiding co-channel interference (CCI). Additionally, with the increasing transmission of sensitive data traffic wirelessly, physical-layer security (PLS) has become a critical concern for a growing number of wireless applications. CSI plays an important role of PLS in numerous applications, such as lightweight user authentication [1] where the channel condition can serve as the signature/fingerprint of a mobile station, and the generation of physical-layer secret keys for data encryption/decryption to enhance secure data transmissions [2], [3].

CSI is, therefore, essential for both enhancing transmission efficiency and ensuring transmission security. In the current state-of-the-art (SOTA) CSI estimation, the Tx (e.g., base station (BS) or access point (AP)) typically broadcasts a predetermined pilot signal, then the Rx (e.g., mobile station) computes the channel coefficients based on the pilot signal it receives and feeds back the estimated CSI to the Tx. Most, if not all, of existing communication systems simply assume that the fed-back CSI is reliable and secure. However, in practice, a malicious terminal may eavesdrop on the feedback link to acquire the legitimate CSI, and then send falsified CSI estimation to the Tx, thus compromising the throughput and/or the security of wireless communications [4], [5], [6], [7], [8], [9], [20], [21], [22]. Although some researchers have already identified the above-mentioned vulnerabilities, demonstrated the feasibility of the CSI-forgery attacks and designed the corresponding countermeasures, they either require modifying the pilot signal at the Tx, implementing pilot encryption operations [4], [5], or necessitate cooperation between the legitimate Rxs and Tx [5], [6], [9], [20], [22], or can merely detect the presence of CSI-forgery without identifying the attackers [4], [5], [6], [8], [20]. As a result, their practical applicability is limited. Furthermore, SOTA studies on CSI-forgery and

Received 7 January 2025; revised 21 August 2025; accepted 20 January 2026. Date of publication 28 January 2026; date of current version 16 February 2026. This work was supported in part by the National Natural Science Foundation of China under Grant 62072351, Grant U23A20300, and Grant 62202359; in part by the Key Research Project of Shaanxi Natural Science Foundation under Grant 2023-JC-ZD-35; in part by the 111 Project under Grant B16037; in part by JSPS KAKENHI under Grant JP25K15087; in part by the Project of Cyber Security Establishment with Inter-University Cooperation; in part by the Science and Technology Research Project of Henan Province under Grant 252102211120; and in part by U.S. National Science Foundation under Grant 2245223. The associate editor coordinating the review of this article and approving it for publication was Prof. Hessam Mahdavifar. *(Corresponding authors: Zhao Li; Jia Liu.)*

Yicheng Liu, Zhao Li, Zheng Yan, and Siwei Le are with the School of Cyber Engineering, Xidian University, Xi'an 710126, China (e-mail: ycliuxdu@stu.xidian.edu.cn; zli@xidian.edu.cn; zyan@xidian.edu.cn; lesiwei6@gmail.com).

Kang G. Shin is with the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109 USA (e-mail: kgshin@umich.edu).

Jia Liu is with the Center for Strategic Cyber Resilience Research and Development, National Institute of Informatics, Tokyo 101-8430, Japan (e-mail: jliu@nii.ac.jp).

Digital Object Identifier 10.1109/TIFS.2026.3659002

countermeasures are usually conducted independently, lacking the exploitation of their commonality. Therefore, it is important to develop a general model for CSI-forgery attacks to facilitate the design of broadly applicable and effective countermeasures.

On the other hand, the development of wireless communication technologies is highly dependent on the advancement of electronic materials and devices. In recent years, reconfigurable intelligent surface (RIS), which is a meta-surface consisting of a large number of passive and controllable reflecting elements [10], has emerged as one of the promising technologies for 6G [11]. Each individual reflecting element in a RIS can be independently controlled with software to modify the amplitude and/or phase shift of the incident signal. This allows the reflected signal to either constructively superimpose with the directly transmitted signal component, thereby enhancing the reception of the desired signal [12], or destructively superimpose with the direct component to nullify the impact of interference or signal at the receiving end, for interference management [13], [14] or security [15], [16], [17], [18]. In essence, due to its low cost, flexibility, and ability to integrate with the wireless environment, one can utilize RIS to tailor the signal propagation environment for specific indoor and outdoor wireless applications. This observation, along with the common practice in most applications where one side of the communication link controls the RIS without the other side's awareness of its use, has led us to leverage RIS to protect a legitimate user's CSI feedback from interception and detect any attempted manipulation of CSI.

Based on the above discussion, we first develop a general model for CSI-forgery attacks that encompasses SOTA methods as special instances, providing a basis for general method design. Then, using this model, we take multi-user multiple-input multiple-output (MU-MIMO) as an example, which is a common communication scenario in both practice and in most existing CSI-based attack research [4], [5], [6], [7], [8], [9], [20], [21], [22], to devise a countermeasure against CSI-forgery attacks. Specifically, we employ a RIS to modify the wireless environment by dynamically adjusting its reflection coefficients matrix. This allows the mobile stations to receive a mixed pilot signal of both the directly transmitted and the reflected components. Therefore, the feedback CSI will differ from that of the direct link between the Tx and mobile stations, preventing the eavesdropping of true CSI. At the Tx-side, by properly configuring the reflection matrix, collecting the CSI feedback from the stations, and analyzing these feedback CSIs, the Tx can detect any manipulated CSI and identify the attacker, thereby ensuring CSI security.

This paper makes the following three main contributions:

- Development of a comprehensive CSI-forgery attack model that can encompass existing CSI-forgery attacks as special instances. Specifically, we introduce a forgery matrix $\mathbf{P}$ to characterize CSI-forgery behaviors. This model provides a foundation to secure CSI feedback.
- Proposal of a RIS-assisted CSI-forgery detection method, called `SecCSI`. By utilizing RIS to dynamically reflect the pilot signal broadcasted from Tx, `SecCSI` can establish a secure environment for preventing CSI interception, detecting the attempts to falsify CSI and identifying the attackers, thus ensuring CSI security and preventing potential malicious subscribers from performing CSI-forgery attacks.
- Experimental validation of the proposed method using the universal software radio peripheral (USRP) platform.

The remainder of this paper is organized as follows. Section II discusses related works on CSI-forgery attacks and Section III describes the system model. Section IV establishes a general model to encompass existing CSI-forgery attacks, while Section V details the design of `SecCSI`. Section VI validates and evaluates the the proposed method, and Section VII concludes the paper.

Throughout this paper, we use the following notations. The set of complex numbers is denoted as $\mathbb{C}$, while vectors and matrices are represented by bold lower-case and upper-case letters, respectively. Let $\mathbf{X}^T$ and $\mathbf{X}^{-1}$ be the transpose and inverse of matrix $\mathbf{X}$. $\|\cdot\|_F$ and $|\cdot|$ indicate the Frobenius norm and the absolute value. $\langle \mathbf{a}, \mathbf{b} \rangle$ represents the inner product of vectors $\mathbf{a}$ and $\mathbf{b}$.

II. Related Work

Accurate and reliable CSI is crucial for improving data transmission efficiency and achieving PLS in communication systems. If the CSI is tampered with by a malicious attacker, and the receiver accepts the manipulated CSI without checking its validity, it could compromise the authentication of legitimate users (LUs) and degrade the transmission performance. Exploiting these vulnerabilities, there have been extensive research on CSI-forgery attacks and their corresponding countermeasures. The authors of [4] proposed for the first time that a forged CSI can be utilized to compromise the efficiency and security of a MU-MIMO system. By eavesdropping on the legitimate CSI and then reporting a forged CSI to the AP, the sniff and power attacks were realized. These attacks can either illegally access the LU's transmission data or steal the AP's transmit power. To counter these threats, the authors devised CSIsec, which enables the AP to broadcast a random training sequence known exclusively to itself, thereby preventing malicious users (MUs) from intercepting the legitimate fed-back CSI and deceiving the AP with a falsified CSI. Following [4], the authors of [5] proposed a polynomial attack to circumvent CSIsec. This attack method relies on the collaboration of several attackers. The authors also suggested the use of time-varying keys to encrypt pilot sequence for channel estimation, known as *Antipoly*, to thwart polynomial attacks. In [6], an analog man-in-the-middle (AMITM) attack was proposed against user authentication based on a wireless link. To prevent AMITM attacks, the Rx can execute additional noise injection (ANI). This noise will be amplified during the attacker's relaying process, making the presence of an AMITM attack detectable. The authors of [7] implemented user selective eavesdropping (USE) in a MU-MIMO system. By generating and reporting a falsified CSI with a gradient variation that preserves orthogonality with the intercepted LU's CSI, both eavesdropping opportunity and effectiveness can be achieved. They proposed *AngleSec*, which leverages channel reciprocity in the angular domain, to combat the USE

TABLE I
COMPARISON OF CSI-FORGERY BASED ATTACKS AND THEIR COUNTERMEASURES

Attacks	Target of attack	Requirement of legitimate CSI	Countermeasure	Pilot modification*	Cooperation of LUs	Identification for MUs
Sniff attack [4]	Confidentiality of legitimate transmission	✓	CSIsec	Data seq.	✗	✗
Polynomial attack [5]		✓	AntiPloy	Data seq.	✓	✗
PCA [9]		✗	Pilot extension	Data seq.	✓	✓
RIS-PSA [21]		✗	Unmentioned	Unmentioned	Unmentioned	Unmentioned
RIS-PCA [22]		✗	GCUSUM	RF sig.	✓	✓
Power attack [4]	Power allocation	✓	CSIsec	Data seq.	✗	✗
AMITM [6]	User scheduling	✓	ANI	✗	✓	✗
USE [7]		✓	AngleSec	✗	✗	✓
MUSTER [8]		✓	RCC	✗	✗	✗
RIS-Jamming [20]	Physical-layer key generation	✗	CPR-CRKG	✗	✓	✗

* The pilot can be modified either by varying the pilot data sequence or by manipulating its radio frequency (RF) waveform. We use "Data seq." and "RF sig." to denote these two methods of pilot modification.

attack. In [8], a MU-MIMO user selection strategy inference and subversion (MUSTER) system was developed to subvert user selection in MU-MIMO networks. By manipulating a falsified CSI with a higher channel gain than that of the LU's CSI, the attacker can seize the service opportunity intended for the LU, thus achieving a denial of service (DoS) attack. Reciprocal consistency checking (RCC), which utilizes the channel reciprocity to identify falsified CSI feedback, was proposed to detect the MUSTER attack. In [9], a pilot contamination attack (PCA) was proposed to degrade the signal-to-noise ratio (SNR) of LUs and simultaneously improve the SNR of MU, by manipulating channel estimation outcomes through malicious pilot insertion. To detect PCA, the authors suggested using a sufficiently long pilot sequence.

As mentioned earlier, RIS has been regarded as one of the promising technologies for 6G [11]. Researchers have demonstrated that RIS can be used for enhancing desired communication [12], managing interference [13], eliminating signal coverage blind spots [14], securing data transmission [15], [16], [17], [18], etc. Essentially, the use of RIS enables researchers to customize the wireless environment for various communication purposes, rather than simply adjusting the transmission to fit the environment. While legitimate Tx/Rx can leverage RIS to enhance the secrecy of communication, malicious attackers may also exploit RIS to compromise legitimate data transmission [19], [20], [21], [22]. Regarding the RIS-assisted CSI-forgery attacks, the authors of [20] investigated the use of RIS in compromising channel reciprocity based key generation (CRKG). By altering the RIS's reflection coefficients, the attacker can disturb the reciprocity of the bi-directional wireless channel, thus sabotaging CRKG process of the legitimate communication pair. To counter the CRKG attack, contaminated path removal based CRKG (CPR-CRKG), which utilizes the auto-correlation coefficient of the path to identify the path created by the RIS, was proposed. Reference [21] utilized RIS to perform pilot spoofing attack (PSA). By adjusting the phase shift at the RIS during the uplink and downlink transmission phases, the channel reciprocity is disrupted, hence biasing the beam towards the eavesdropper. The authors of [22] leveraged RIS to implement pilot contamination attack (PCA), known as RIS-PCA, where the attacker employs a RIS to reflect the pilot signal from the LU in the uplink training phase. To detect RIS-PCA, a

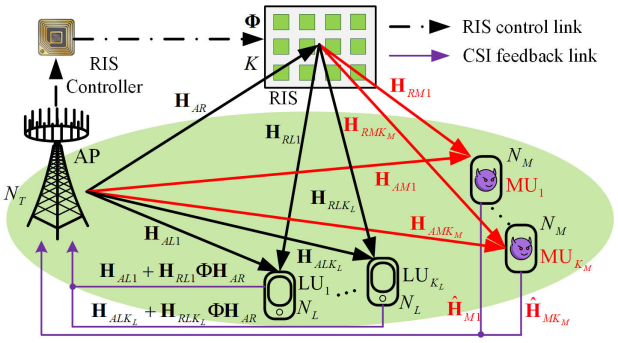


Fig. 1. System model.

generalized cumulative sum (GCUSUM) was proposed, which can detect RIS-PCA attacks by comparing the log-likelihood ratio of a normally received signal with that of a signal received under attack.

We compare existing CSI-forgery attacks and their countermeasures in Table I. We can conclude from the above literature review that existing countermeasures are typically designed for particular types of attack, thus leaving them vulnerable to other types of attack. Some of them (1) rely on modification of the pilot data sequence, necessitating protocol changes and limiting their applicability, or (2) require the cooperation between the legitimate Rxs and Tx, incurring additional overhead at the Rxs, or (3) can only identify the occurrence of CSI-forgery without attributing it to a specific attacker. These deficiencies mainly come from the fact that SOTA approaches aim to function as specific patches in unreliable environments to achieve PLS. It is, therefore, important to develop an approach to mitigate these deficiencies. Instead of designing new patches, in contrast, our objective is to establish a secure environment to proactively prevent any potential attacks based on CSI-forgery.

III. SYSTEM MODEL

We consider a MU-MIMO downlink communication system, consisting of an AP equipped with N_T antennas, K_L legitimate users (LUs) each with N_L antennas, and K_M malicious users (MUs) each with N_M antennas, as illustrated in Fig. 1. For simplicity, we let $N_L = N_M = N_R$.

The AP controls a RIS composed of K reflecting elements via a RIS controller. We use $\mathbf{H}_{ALl} \in \mathbb{C}^{N_R \times N_T}$ and $\mathbf{H}_{AMm} \in$

$\mathbb{C}^{N_R \times N_T}$ where $l \in \{1, \cdots, K_L\}$ and $m \in \{1, \cdots, K_M\}$, to denote the channel matrices of the direct links from the AP to the lth LU and the mth MU. Additionally, $\mathbf{H}_{AR} \in \mathbb{C}^{K \times N_T}$, $\mathbf{H}_{RLl} \in \mathbb{C}^{N_R \times K}$, and $\mathbf{H}_{RMm} \in \mathbb{C}^{N_R \times K}$ represent the reflecting links from the AP to RIS, RIS to LU$_l$, and RIS to MU$_m$, respectively. The distances from the RIS to the AP and the users are sufficiently large so that the reflecting and direct links are independent from each other. We adopt a spatially uncorrelated Rayleigh flat fading channel to model the elements of the above channel matrices as independent and identically distributed (i.i.d.) zero-mean unit-variance complex Gaussian random variables. We assume that the wireless environment exhibits quasi-static/block-fading characteristics. Considering our focus on guaranteeing the truthfulness of the CSI feedback from the users, we assume that AP can accurately acquire $\mathbf{H}_{AR} \in \mathbb{C}^{K \times N_T}$ [23]. Additionally, all LUs and MUs feed their estimated or forged CSI back to the AP [24]. The delay of CSI delivery is smaller than the coherent time, within which the wireless channel remains constant [25]. The MU can eavesdrop on the CSI feedback from the target LU to the AP [4], [5], [6], [7]. Furthermore, upon acquiring the CSI of the target LU, the MU can fabricate CSI and send the falsified CSI back to the AP with the intention of either disrupting the legitimate data transmission [4], [6], [7], [8], [20] or gaining unauthorized access to the legitimate information [4], [5], [9], [21], [22]. We use diagonal matrix $\mathbf{\Phi} = \text{diag}\{[\phi_1, \phi_2, \cdots, \phi_K]\} \in \mathbb{C}^{K \times K}$ to denote the reflection matrix of the RIS. $\phi_k = \eta_k e^{j\theta_k}$ ($k \in \{1, \cdots, K\}$) represents the reflection coefficient of the kth element of the RIS, where $\eta_k \in [0, 1]$ and $\theta_k \in [0, 2\pi]$ are the absorption coefficient and the phase shift coefficient of the element, respectively. Since the environment can involve multipath, the use of the RIS remains transparent to both LUs and MUs. We assume that the RIS is securely controlled by the AP via a RIS controller [16], [17], [26]. Consequently, the MUs are unable to detect the presence of the RIS or its reflecting coefficients. Additionally, `SecCSI` does not modify the pilot sequence; this information is publicly known and thus accessible to potential attackers.

IV. General Model for CSI-Forgery Attacks

We propose a general model that can subsume existing CSI-forgery attacks as special cases. This model can be used to facilitate the development of general countermeasures against CSI-forgery attacks. Like most related works [4], [5], [6], [7], [8], [9], [20], [21], [22], we use MU-MIMO downlink system as an example to formulate our general model for CSI-forgery attacks. The estimated CSI of lth LU and mth MU can be represented as $\mathbf{H}_{Ll} \in \mathbb{C}^{N_R \times N_T}$ and $\mathbf{H}_{Mm} \in \mathbb{C}^{N_R \times N_T}$, where $l \in \{1, \cdots, K_L\}$ and $m \in \{1, \cdots, K_M\}$, respectively. We define a combined CSI matrix consisting of κ_L ($1 \le \kappa_L \le K_L$) LUs' and κ_M ($1 \le \kappa_M \le K_M$) MUs' estimations as $\mathbf{H} = [\mathbf{H}_{L1}^T \ \cdots \ \mathbf{H}_{M1}^T \ \cdots \ \mathbf{H}_{L\kappa_L}^T \ \cdots \ \mathbf{H}_{M\kappa_M}^T]^T$.[1]

[1]Since the order of feedback can be random in practice, we use ellipses to indicate that falsified CSI can appear at any position in $\mathbf{H}$, without loss of generality. Specifically, the ellipsis between $\mathbf{H}_{L1}^T$ and $\mathbf{H}_{M1}^T$ signifies that one or more genuine CSIs reported by a single LU or multiple LUs may appear between $\mathbf{H}_{L1}^T$ and $\mathbf{H}_{M1}^T$.

So, the AP will perceive a combined feedback CSI matrix $\mathbf{F} = [\mathbf{F}_{L1}^T \ \cdots \ \mathbf{F}_{M1}^T \ \cdots \ \mathbf{F}_{L\kappa_L}^T \ \cdots \ \mathbf{F}_{M\kappa_M}^T]^T$.

As mentioned earlier, there may be multiple MUs who provide falsified CSI feedback, and thus $\mathbf{F}$ may not be equal to $\mathbf{H}$. To characterize such a situation, we introduce a forgery matrix $\mathbf{P} \in \mathbb{C}^{\kappa N_R \times \kappa N_R}$ to establish the relationship between $\mathbf{H}$ and $\mathbf{F}$, as given by Eq. (1).

$$\mathbf{F} = \mathbf{P}\mathbf{H}. \tag{1}$$

We can easily observe from the above equation that when all of the mobile users are honest, $\mathbf{P}$ becomes an identity matrix and $\mathbf{F} = \mathbf{H}$ holds. Then, as long as a MU indexed with m ($m \in \{1, \cdots, \kappa_M\}$) feeds back a forged CSI, denoted as $\hat{\mathbf{H}}_{Mm}$, instead of the estimated true $\mathbf{H}_{Mm}$, to the AP, we can have $\mathbf{F} = [\mathbf{H}_{L1}^T \ \cdots \ \mathbf{H}_{M1}^T \ \cdots \ \hat{\mathbf{H}}_{Mm}^T \ \cdots \ \mathbf{H}_{L\kappa_L}^T \ \cdots \ \mathbf{H}_{M\kappa_M}^T]^T$, leading to $\mathbf{F} \neq \mathbf{H}$. Based on this observation, existing CSI-forgery behaviors can be represented by various constructions of $\mathbf{P}$.

In what follows, we will use three typical CSI-forgery attacks — namely, the sniff attack [4], USE attack [7], and RIS-Jamming attack [20] — as examples to demonstrate that our general model can encompass these CSI-forgery attacks as special instances. Other attacks can be analyzed similarly. Due to space limitations, we omit redundant explanations in this paper. For clarity of exposition, we illustrate the aforementioned three types of attacks in Fig. 2.

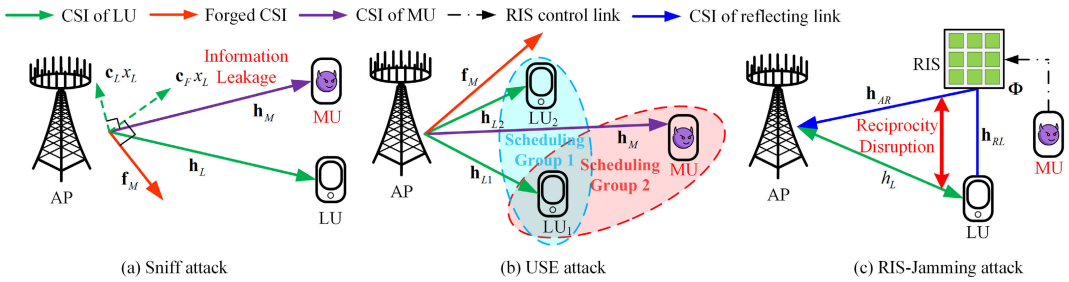
- **Sniff attack:** The authors of [4] set $K_L = K_M = 1$, $N_T = 2$, and $N_L = N_M = 1$, and considered a downlink communication system where the AP employs zero-forcing beamforming (ZFBF) to pre-process multiple data streams for its serving subscribers. In this configuration, channel matrices $\mathbf{H}_{Ll}$ and $\mathbf{H}_{Mm}$ become vectors $\mathbf{h}_L \in \mathbb{C}^{1\times 2}$ and $\mathbf{h}_M \in \mathbb{C}^{1\times 2}$. The AP utilizes precoding matrix $\mathbf{H}^{-1}$ where $\mathbf{H} = [\mathbf{h}_L^T \ \mathbf{h}_M^T]^T$ to process the data x_L and x_M for transmission to the LU and MU, respectively. However, when the MU provides a falsified CSI $\hat{\mathbf{h}}_M = [\hat{h}_M^{(1)} \ \hat{h}_M^{(2)}]$ instead of its genuine CSI $\mathbf{h}_M$, the combined feedback CSI at the AP becomes $\mathbf{F} = [\mathbf{h}_L^T \ \hat{\mathbf{h}}_M^T]^T$. As a result, the AP will utilize $\mathbf{F}^{-1}$ as the precoder, and the reception of this system can be expressed as:

$$\begin{aligned}\begin{bmatrix} y_L \\ y_M \end{bmatrix} &= \mathbf{H}\mathbf{F}^{-1} \begin{bmatrix} \sqrt{P_L}x_L \\ \sqrt{P_M}x_M \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ w_1 & w_2 \end{bmatrix} \begin{bmatrix} \sqrt{P_L}x_L \\ \sqrt{P_M}x_M \end{bmatrix},\end{aligned} \tag{2}$$

where y_L and y_M represent the received signals at the LU and MU, P_L and P_M are the transmit power allocated for transmissions to the LU and MU, x_L and x_M are the data symbols sent for the LU and MU, respectively. By calculating $\mathbf{H}\mathbf{F}^{-1}$ to obtain w_1 and w_2, the Sniff attack can be characterized as:

$$\begin{aligned}\mathbf{P} &= \begin{bmatrix} 1 & 0 \\ w_1 & w_2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ \frac{h_M^{(1)}\hat{h}_M^{(2)} - h_M^{(2)}\hat{h}_M^{(1)}}{h_L^{(1)}\hat{h}_M^{(2)} - \hat{h}_M^{(1)}h_L^{(2)}} & \frac{h_L^{(2)}h_M^{(2)} - h_L^{(2)}h_M^{(1)}}{h_L^{(1)}\hat{h}_M^{(2)} - \hat{h}_M^{(1)}h_L^{(2)}} \end{bmatrix}^{-1}.\end{aligned} \tag{3}$$

Therefore, the MU can manipulate w_2 to approach 0 while maximizing w_1 by intercepting the legitimate $\mathbf{h}_L$ and providing a falsified $\hat{\mathbf{h}}_M$ to the AP.



(a) Sniff attack (b) USE attack (c) RIS-Jamming attack

Fig. 2. Three typical CSI-forgery attacks.

• **USE attack:** The authors of [7] implemented USE attack in a MU-MIMO system consisting of one AP equipped with two antennas, two single-antenna LUs, and one single-antenna MU. The two LUs are denoted as LU_1 and LU_2. The CSI between the AP and mobile users can be represented as $\mathbf{h}_{L1} = [h_{L1}^{(1)}\ h_{L1}^{(2)}]$, $\mathbf{h}_{L2} = [h_{L2}^{(1)}\ h_{L2}^{(2)}]$, and $\mathbf{h}_M = [h_M^{(1)}\ h_M^{(2)}]$, respectively. In this configuration, the AP can only serve two out of the three users simultaneously in a transmission slot. To maximize system throughput, the AP may select the two users with uncorrelated or highly orthogonal channel characteristics. In order to seize the scheduling opportunity, the MU can provide a forged CSI, $\hat{\mathbf{h}}_M$, to the AP.

Without loss of generality, we take the MU seizing scheduling opportunity from LU_2 as an example. Consequently, the combined CSI at the AP becomes $\mathbf{F} = [\mathbf{h}_{L1}^T\ \mathbf{h}_{L2}^T\ \hat{\mathbf{h}}_M^T]^T$ instead of $\mathbf{F} = \mathbf{H} = [\mathbf{h}_{L1}^T\ \mathbf{h}_{L2}^T\ \mathbf{h}_M^T]^T$. Since $\langle \mathbf{h}_{L1}, \mathbf{h}_{L2} \rangle \neq 0$ always holds in practice, then by ensuring $\langle \mathbf{h}_{L1}, \hat{\mathbf{h}}_M \rangle = 0$, this falsified $\hat{\mathbf{h}}_M$ can facilitate the MU to be scheduled with LU_1.

According to our CSI-forgery model, the USE attack can be characterized by:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ w_1 & w_2 & w_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\alpha \frac{h_{L2}^{(1)} h_{L1}^{(1)} + h_{L2}^{(2)} h_{L1}^{(2)}}{(h_{L1}^{(1)})^2 + (h_{L1}^{(2)})^2} & -\beta \frac{\hat{h}_M^{(1)} h_{L1}^{(1)} + \hat{h}_M^{(2)} h_{L1}^{(2)}}{(h_{L1}^{(1)})^2 + (h_{L1}^{(2)})^2} & \alpha\ \beta \end{bmatrix}, \tag{4}$$

where α and β are arbitrary complex numbers, satisfying $\langle \mathbf{h}_{L1}, \hat{\mathbf{h}}_M \rangle = 0$.

• **RIS-Jamming attack:** The authors of [20] implemented the RIS-Jamming attack to disrupt the channel reciprocity between the AP and the LU. They set $K_L = K_M = 1$, $N_T = 1$, $N_L = N_M = 1$, and $K \geq 1$ in their design, resulting in $\mathbf{H}_{AR}$ and $\mathbf{H}_{RL}$ becoming column vector $\mathbf{h}_{AR}$ and row vector $\mathbf{h}_{RL}$, respectively. To implement the RIS-Jamming attack, the MU prompts the AP to perceive a falsified feedback CSI as:

$$f_{LUp} = h_{LUp} + \mathbf{h}_{AR}^T \mathbf{\Phi} \mathbf{h}_{RL}^T. \tag{5}$$

We define h_{LUp} and h_{LDown} as the uplink and downlink CSI without RIS intervention, and $h_{LUp} = h_{LDown}$ holds. To mount RIS-Jamming attack, the MU adjusts $\mathbf{\Phi}$ to produce $f_{LUp} \neq h_{LDown}$ where $f_{LUp} = p_M h_{LDown}$. Here p_M characterizes the MU's CSI-forgery behavior, which is derived from $\mathbf{P}$ under $N_L = N_M = 1$ and single LU feedback configurations. Therefore, we can model the RIS-Jamming attack as $p_M = 1 + \frac{\mathbf{h}_{AR}^T \mathbf{\Phi} \mathbf{h}_{RL}^T}{h_{LUp}}$, which is consistent with our general model.

From the above discussion about the three representative CSI-forgery attacks, it is evident that although these attacks are distinct from each other, they share commonalities that allow us to establish a general CSI-forgery model using a matrix $\mathbf{P}$ to capture attacker behaviors. This model not only encompasses existing CSI-forgery attacks as special cases but also generalizes to potential unknown ones, providing the foundation for designing a comprehensive prevention strategy against this type of threat.

V. Design of SecCSI

From our earlier findings, it is evident that when $\mathbf{P} = \mathbf{I}$, i.e., $\mathbf{P}$ is an identity matrix, the feedback CSI remains secure. This motivates us to develop a detection method that aims to: 1) scrutinize $\mathbf{P}$ to determine the presence of CSI-forgery, and 2) identify the MU responsible for providing a falsified CSI. To achieve this goal, we will employ RIS to establish a secure wireless environment, referred to as SecCSI, for CSI-forgery detection and attacker identification.

A. Detection of CSI-Forgery

We formulate the actual CSI between the AP and $K_L + K_M$ users as $\mathbf{H}_R \mathbf{\Phi} \mathbf{H}_{AR} + \mathbf{H}_D$, where $\mathbf{\Phi}$ is the reflecting matrix of the RIS. We define two combined CSI matrices $\mathbf{H}_R = [\mathbf{H}_{RL1}^T \cdots \mathbf{H}_{RM1}^T \cdots \mathbf{H}_{RMK_M}^T \cdots \mathbf{H}_{RLK_L}^T]^T \in \mathbb{C}^{(K_L+K_M)N_R \times K}$ and $\mathbf{H}_D = [\mathbf{H}_{AL1}^T \cdots \mathbf{H}_{AM1}^T \cdots \mathbf{H}_{AMK_M}^T \cdots \mathbf{H}_{ALK_L}^T]^T \in \mathbb{C}^{(K_L+K_M)N_R \times N_T}$ to indicate the CSI of the reflecting link from the RIS to users and the direct link from the AP to users, respectively. For clarity of presentation, we begin with $K_L = K_M = 1$. Therefore, we can obtain $\mathbf{H}_R = [\mathbf{H}_{RL}^T\ \mathbf{H}_{RM}^T]^T \in \mathbb{C}^{2N_R \times K}$ and $\mathbf{H}_D = [\mathbf{H}_{AL}^T\ \mathbf{H}_{AM}^T]^T \in \mathbb{C}^{2N_R \times N_T}$. Since we consider only one LU and MU, we can omit the indices of LU and MU without causing ambiguity.

In the CSI feedback stage, the LU feeds back $\mathbf{H}_L = \mathbf{H}_{RL} \mathbf{\Phi} \mathbf{H}_{AR} + \mathbf{H}_{AL}$, while the MU provides a forged CSI, $\hat{\mathbf{H}}_M$, instead of its genuine CSI, $\mathbf{H}_M = \mathbf{H}_{RM} \mathbf{\Phi} \mathbf{H}_{AR} + \mathbf{H}_{AM}$, to the AP. Then, according to Eq. (1), the combined feedback CSI at the AP can be expressed as:

$$\mathbf{F} = \begin{bmatrix} \mathbf{H}_L \\ \hat{\mathbf{H}}_M \end{bmatrix} = \mathbf{P} \begin{bmatrix} \mathbf{H}_L \\ \mathbf{H}_M \end{bmatrix} = \mathbf{P} \left(\mathbf{H}_R \mathbf{\Phi} \mathbf{H}_{AR} + \mathbf{H}_D \right). \tag{6}$$

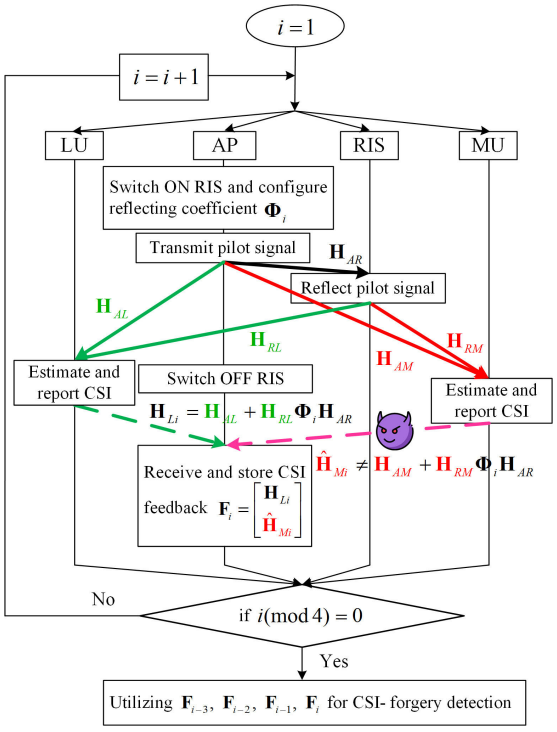


Fig. 3. Illustration of pilot manipulation flow in SecCSI.

The process of SecCSI involves four rounds of channel estimation and feedback (CEF) in each block,[2] where we set $\mathbf{\Phi}$ to $\mathbf{\Phi}_i$, $\mathbf{H}_L$ to $\mathbf{H}_{Li}$, and $\mathbf{H}_M$ to $\mathbf{H}_{Mi}$, respectively, with $i \in \{1, 2, 3, 4\}$. We use Fig. 3 to illustrate the pilot manipulation flow in SecCSI. For simplicity, we consider $K_L = K_M = 1$ in this example.

As the figure shows, during each CEF round, the AP dynamically configures the ON/OFF states and reflecting coefficients of the RIS. Specifically, the AP first activates the RIS and broadcasts a pilot signal, which reaches both the LU and the MU via the direct and reflecting links. After transmitting the pilot signal, the AP deactivates the RIS. Meanwhile, the users estimate their CSIs as $\mathbf{H}_{Li} = \mathbf{H}_{AL}\mathbf{\Phi}_i\mathbf{H}_{AR}$ and $\mathbf{H}_{Mi} = \mathbf{H}_{AM}\mathbf{\Phi}_i\mathbf{H}_{AR}$. Since the received signal at the user is a superposition of direct and reflected components, the MU cannot extract the genuine CSI of its direct link (i.e., $\mathbf{H}_{AM}$), effectively achieving CSI concealment. Subsequently, the LU and MU report genuine CSI ($\mathbf{H}_{Li}$) and falsified CSI ($\hat{\mathbf{H}}_{Mi}$), respectively, to the AP, where the feedback CSI matrix is combined as $\mathbf{F}_i = [\mathbf{H}_{Li}^T \ \mathbf{H}_{Mi}^T]^T$. Then, any CSI-forgery can be detected by analyzing the feedback from four consecutive CEF rounds, denoted as $\mathbf{F}_{i-3}, \mathbf{F}_{i-2}, \mathbf{F}_{i-1}$, and $\mathbf{F}_i$, where i must selected such that $i \bmod 4 = 0$ holds.

It is important to note that SecCSI requires four rounds of CEF to effectively detect CSI-forgery attacks. This detection process is achieved by comparing successive CSI feedback from users. We provide Proposition 1 to clarify this requirement. Its proof can be found in Appendix A.

Proposition 1: SecCSI requires at least four rounds of CEF to detect CSI-forgery attacks.

Taking the first two rounds of CEF as an example, we can obtain the combined feedback CSI at the AP as:

$$\mathbf{F}_i = \mathbf{P}_i \left(\mathbf{H}_R \mathbf{\Phi}_i \mathbf{H}_{AR} + \mathbf{H}_D\right), \tag{7}$$

where $\mathbf{F}_i$ and $\mathbf{P}_i$ ($i \in \{1, 2\}$) represent the combined feedback CSI at the AP and the feedback behavior matrix of all users in the ith round of CEF, respectively. It is important to note that since $\mathbf{H}_R$, $\mathbf{H}_{AR}$, and $\mathbf{H}_D$ are genuine CSIs, whereas the CSI-forgery behavior is exclusively modeled by $\mathbf{P}_i$, we delete the subscripts i of these matrices for simplicity.

We use $\phi_{ki} = \eta_{ki} e^{j\theta_{ki}}$, where $k \in \{1, \cdots, K\}$, to denote the kth reflecting coefficient of $\mathbf{\Phi}_i$ in the ith round of CEF. For simplicity, we configure all of the K reflecting coefficients to be the same in a single CEF. Then, by setting $\theta_{k1}, \theta_{k3} \in [0, \pi]$, satisfying $\theta_{k1} \neq \theta_{k3}$, $\theta_{k2} = \theta_{k1} + \pi$, and $\theta_{k4} = \theta_{k3} + \pi$, we can have $\mathbf{\Phi}_2 = -\mathbf{\Phi}_1$. Substituting $\mathbf{\Phi}_2 = -\mathbf{\Phi}_1$ into Eq. (7) and subtracting $\mathbf{F}_2$ from $\mathbf{F}_1$, we can get:

$$\mathbf{F}_1 - \mathbf{F}_2 = (\mathbf{P}_1 + \mathbf{P}_2)(\mathbf{H}_R \mathbf{\Phi}_1 \mathbf{H}_{AR}) + (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{H}_D. \tag{8}$$

When there is no CSI-forgery during the first two rounds of CSI estimation, $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{I}$ can hold.[3] Moreover, since both $\mathbf{\Phi}_1$ and $\mathbf{H}_{AR}$ are available at the AP [23] and are immune to CSI-forgery, we can derive $\mathbf{H}_R$ from Eq. (8) as:

$$\mathbf{H}_R = \frac{1}{2}(\mathbf{F}_1 - \mathbf{F}_2)(\mathbf{\Phi}_1 \mathbf{H}_{AR})^{-1}. \tag{9}$$

It is worth noting that $\mathbf{\Phi}_1\mathbf{H}_{AR}$ must be a square matrix in order to achieve $(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1}$. This requirement can be easily satisfied because the number of reflecting elements on the RIS is always greater than the number of transmit antennas at the AP, i.e., $K > N_T$ holds. Therefore, by activating an appropriate number of reflecting elements on the RIS, $K = N_T$ can be ensured, making $\mathbf{\Phi}_1\mathbf{H}_{AR}$ a square matrix.

Similar to the derivation of Eq. (9), without CSI-forgery, we can derive $\mathbf{H}_R$ based on the 3rd and 4th rounds of CEF as $\mathbf{H}_R = \frac{1}{2}(\mathbf{F}_3 - \mathbf{F}_4)(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1}$. By this equation and Eq. (8),

[2]The length of a block depends on the channel coherent time [27]. In SecCSI, we consider a commonly used quasi-static channel model, where the channel coherent time typically ranges from one hundred to several hundreds of milliseconds [28]. Since the the number of users has a dominant influence on the total time overhead of CEF, as each user must sequentially report its CSI back to the AP after receiving the broadcast pilot signal and performing channel estimation. Consequently, the time overhead for one round of CEF can vary from several milliseconds to several tens of milliseconds, depending on the network scale and user density [29]. Therefore, SecCSI can easily accommodate four rounds of CEF within the channel coherent time when the number of users is not excessively high. However, it is important to note that when SecCSI is applied to a system with shorter channel coherent time or a large number of users, we can employ user scheduling methods to divide the users into multiple smaller groups. This approach allows SecCSI to be applied directly to each group, making the method scalable.

[3]Here, we assume that no CSI-forgery occurs during the four rounds of the CEF, meaning that $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{I}$ holds. Based on this assumption, we derive Eqs. (8)–(10) to establish the conditions under which CSI-forgery is absent. It is important to note that we impose no restrictions on the behavior of CSI-forgery. This aspect will be explored through an analysis of 15 different types of CSI-forgery behaviors. Furthermore, there may be scenarios in which a user fails to report CSI in one or more rounds of CEF. In such cases, the AP will automatically classify that user as a MU. That is, providing CSI feedback in the four rounds of CEF is a prerequisite for subsequent data transmission.

without CSI-forgery during the four successive rounds of CEF, the following equality holds.

$$(\mathbf{F}_1 - \mathbf{F}_2)(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1} - (\mathbf{F}_3 - \mathbf{F}_4)(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1} = \mathbf{0}. \quad (10)$$

Based on the above analysis, by checking the validity of Eq. (10) using the four $\mathbf{F}_i$s where $i \in \{1, 2, 3, 4\}$, the AP can detect the presence of CSI-forgery. When the AP verifies that no CSI-forgery has occurred, it can simply calculate $\frac{1}{2}(\mathbf{F}_1+\mathbf{F}_2)$ or $\frac{1}{2}(\mathbf{F}_3+\mathbf{F}_4)$ to yield $\mathbf{H}_D$. Therefore, the use of RIS does not affect the AP's acquisition of the actual CSI.

In practice, the MU may provide falsified CSI to the AP in one or more rounds of CEF. Below we will elaborate on the CSI-forgery detection method while taking into account various CSI-forgery behaviors. We first provide Eq. (11) in terms of Eq. (8) as:

$$\mathbf{F}_3 - \mathbf{F}_4 = (\mathbf{P}_3 + \mathbf{P}_4)(\mathbf{H}_R\mathbf{\Phi}_3\mathbf{H}_{AR}) + (\mathbf{P}_3 - \mathbf{P}_4)\mathbf{H}_D. \quad (11)$$

By substituting Eqs. (8) and (11) into Eq. (10), we can derive Eq. (12), as shown at the bottom of the page, where $\mathbf{0}$ denotes zero matrix. Eq. (12) is equivalent to Eq. (10). When $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{I}$, Eq. (12) holds true, indicating that all users are honest. Otherwise, when $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ take on various values, representing different CSI-forgery behaviors, Eq. (12) does not hold, implying occurrence of CSI-forgery.

As can be easily inferred, during the four rounds of CEF, there are a total of 15 modes of CSI-forgery behaviors: 1) $\mathbf{P}_1 \neq \mathbf{P}_2 \neq \mathbf{P}_3 \neq \mathbf{P}_4$, 2) $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 \neq \mathbf{P}_4$, 3) $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_4 \neq \mathbf{P}_3$, 4) $\mathbf{P}_1 = \mathbf{P}_3 = \mathbf{P}_4 \neq \mathbf{P}_2$, 5) $\mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 \neq \mathbf{P}_1$, 6) $\mathbf{P}_1 = \mathbf{P}_2 \neq \mathbf{P}_3 = \mathbf{P}_4$, 7) $\mathbf{P}_1 = \mathbf{P}_3 \neq \mathbf{P}_2 = \mathbf{P}_4$, 8) $\mathbf{P}_1 = \mathbf{P}_4 \neq \mathbf{P}_2 = \mathbf{P}_3$, 9) $\mathbf{P}_1 = \mathbf{P}_2 \neq \mathbf{P}_3 \neq \mathbf{P}_4$, 10) $\mathbf{P}_1 = \mathbf{P}_3 \neq \mathbf{P}_2 \neq \mathbf{P}_4$, 11) $\mathbf{P}_1 = \mathbf{P}_4 \neq \mathbf{P}_2 \neq \mathbf{P}_3$, 12) $\mathbf{P}_2 = \mathbf{P}_3 \neq \mathbf{P}_1 \neq \mathbf{P}_4$, 13) $\mathbf{P}_2 = \mathbf{P}_4 \neq \mathbf{P}_1 \neq \mathbf{P}_3$, 14) $\mathbf{P}_3 = \mathbf{P}_4 \neq \mathbf{P}_1 \neq \mathbf{P}_2$, and 15) $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 \neq \mathbf{I}$.

By observing Eq. (12), since $\mathbf{H}_R$, $\mathbf{H}_D$, and $\mathbf{H}_{AR}$ are random matrices, it is apparent that Eq. (12) will not hold as long as $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ are not identical. Therefore, CSI-forgery modes 1 to 14 can be detected when Eq. (12) (or equivalently Eq. (10)) does not hold.[4] As for the CSI-forgery mode indexed with 15, Eq. (12) holds true, so the criterion defined by Eq. (10) is no longer applicable. To tackle this particular CSI-forgery mode, we design a supplementary detection approach to enhance SecCSI. To assist in this supplementary method design, we introduce a non-zero matrix $\mathbf{Q}_i$ ($i \in \{1, 2, 3, 4\}$), which can also characterize the CSI-forgery behavior as $\mathbf{P}_i$ does, such that:

$$\mathbf{F}_i = \mathbf{P}_i\mathbf{H} = \mathbf{H}\mathbf{Q}_i. \quad (13)$$

According to Eq. (13), when $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_s \neq \mathbf{I}$, a $\mathbf{Q}_s$ can be derived similarly. Subsequently, by substituting Eq. (13) into Eqs. (8) and (11), and then left-multiplying $(\mathbf{F}_3 - \mathbf{F}_4)^{-1}$ by $\mathbf{F}_1 - \mathbf{F}_2$, we can have[5]:

$$(\mathbf{F}_3 - \mathbf{F}_4)^{-1}(\mathbf{F}_1 - \mathbf{F}_2) = \mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s. \quad (14)$$

Here we provide Proposition 2, according to which we can derive Eq. (15) from Eq. (14). The proof of Proposition 2 can be found in Appendix B.

Proposition 2: If and only if $\mathbf{Q}_s = \alpha\mathbf{I}$ where α is a non-zero scalar, the equation $\mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s = \mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}$ holds.

$$(\mathbf{F}_3 - \mathbf{F}_4)^{-1}(\mathbf{F}_1 - \mathbf{F}_2) - \mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR} = \mathbf{0}. \quad (15)$$

Since $\mathbf{\Phi}_3$ is a diagonal matrix, and we can ensure $\mathbf{H}_R$ and $\mathbf{H}_{AR}$ to be square matrices by arranging the feedback CSIs at the AP to meet $\kappa N_R = K$ and controlling the number of activated elements on the RIS to meet $K = N_T$, respectively, their inverses exist. Additionally, given that $\mathbf{F}_i$ where $i \in \{1, 2, 3, 4\}$, $\mathbf{H}_{AR}$, $\mathbf{\Phi}_1$, and $\mathbf{\Phi}_3$ are available at the AP, the AP can compute both $(\mathbf{F}_3 - \mathbf{F}_4)^{-1}(\mathbf{F}_1 - \mathbf{F}_2)$ and $\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}$. Therefore, SecCSI can detect the CSI-forgery mode 15. In practice, we can first apply Eq. (10) to detect the occurrence of CSI-forgery modes other than $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 \neq \mathbf{I}$. If no CSI-forgery is detected, we can proceed to apply Eq. (15) to ascertain whether the MU consistently presents the same falsified CSI feedback across the four rounds of CEF.

So far, we have presented the design of RIS-assisted CSI-forgery detection under $K_L = K_M = 1$. However, our method is applicable even when $K_L > 1$ and $K_M > 1$ because under SecCSI, the AP sets $K = N_T$ to ensure $\mathbf{\Phi}_i\mathbf{H}_{AR}$ ($i \in \{1, 3\}$) is a square matrix, and groups $\kappa = N_T/N_R$ users' CSI feedback to ensure $\mathbf{F}_i$ is also a square matrix for CSI-forgery detection. Then, as long as CSI-forgery occurs, regardless of the number of falsified CSIs fed to the AP, it can be easily seen that Eq. (10) will not hold true for CSI-forgery modes 1 to 14. Otherwise, Eq. (15) will not hold for CSI-forgery mode 15. This way, a comprehensive CSI-forgery detection is achieved.

B. Identification for MUs

Upon detecting a CSI-forgery, SecCSI will further identify the MUs who provide falsified CSIs to the AP. For clarity of exposition on MUs' identification, we begin with setting $K_L > 1$ and $K_M = 1$, satisfying $(K_L + K_M)N_R > N_T > N_R$.

[4]Although specific non-identity matrices $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ may satisfy Eq. (12), given the large number of elements in the random CSI matrices, accurately determining non-identical $\mathbf{P}_1$, $\mathbf{P}_2$, $\mathbf{P}_3$, and $\mathbf{P}_4$ to satisfy Eq. (12) can be considered impossible in practice.

[5]It is worth noting that $\mathbf{F}_3 - \mathbf{F}_4 \in \mathbb{C}^{(K_L+K_M)N_R \times N_T}$ must be a square matrix in order to ensure the existence of its inverse. This requirement can be easily met because the values of N_T and N_R are typically powers of 2 in practical use. Additionally, during the CEF process, multiple users provide their CSI feedback to the AP sequentially. These characteristics enable the AP to arrange $\mathbf{F}_i$ as a square matrix. Specifically, during the CSI feedback stage, the AP constructs $\mathbf{F}_i$ based on the reports from the users. For each group of $\kappa \leq K_L + K_M$ feedback CSIs satisfying $\kappa N_R = N_T$, $\mathbf{F}_i$ is a square matrix. In other words, our method first divides $K_L + K_M$ users into $\lceil (K_L + K_M)/\kappa \rceil$ groups, where $\lceil \cdot \rceil$ represents rounding up to the nearest integer, and then detects CSI-forgery in each group. When $(K_L + K_M)$ is not divisible by κ, we can group the last κ feedback CSIs together for detection. However, this approach incurs the cost of verifying $\kappa - (K_L + K_M)$ mod κ CSIs twice where mod denotes taking the remainder.

$$[(\mathbf{P}_1 + \mathbf{P}_2) - (\mathbf{P}_3 + \mathbf{P}_4)]\mathbf{H}_R + (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{H}_D(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1} - (\mathbf{P}_3 - \mathbf{P}_4)\mathbf{H}_D(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1} = \mathbf{0}. \quad (12)$$

Therefore, during each round of CEF, the AP needs to organize each $\kappa = N_T/N_R$ feedback CSIs as an identification group. There will be $G = \lceil (K_L + K_M)/\kappa \rceil$ groups in total for parallel identification.

For the identification group involving falsified CSI, say $\mathbf{F}_i^0$ (the superscript 0 indicates that the matrix has not yet been divided), we first divide it into two sub-groups, denoted as $\mathbf{F}_i^{11}$ and $\mathbf{F}_i^{12}$ and combine each with matrix $\mathbf{T}_1$ consisting of $N_T/(2N_R)$ reliable CSIs,[6] to form $[(\mathbf{F}_i^{11})^T \ \mathbf{T}_1^T]^T$ and $[(\mathbf{F}_i^{12})^T \ \mathbf{T}_1^T]^T$. These combinations can then be further inspected using Eqs. (10) and (15). Without loss of generality, we assume that the CSI-forgery occurs in $\mathbf{F}_i^{12}$. Then, $[(\mathbf{F}_i^{11})^T \ \mathbf{T}_1^T]^T$ will be verified to be not falsified, and $\mathbf{F}_i^{12}$ will be divided into two sub-groups, each containing $N_T/(4N_R)$ feedback CSIs, denoted as $\mathbf{F}_i^{21}$ and $\mathbf{F}_i^{22}$, respectively. The AP then inspects $[(\mathbf{F}_i^{21})^T \ \mathbf{T}_2^T]^T$ and $[(\mathbf{F}_i^{22})^T \ \mathbf{T}_2^T]^T$ where $\mathbf{T}_2$ consists of $3N_T/(4N_R)$ reliable CSIs. We repeat the above process until the MU that provides falsified CSI is identified in either $\mathbf{F}_i^{D1}$ or $\mathbf{F}_i^{D2}$, where $D = \log_2(N_T/N_R)$ represents the maximum time of MU identification. Although the above design is based on the single-MU assumption, this identification method can be directly applied even when there are multiple MUs.

C. SecCSI's Effectiveness Against Collaborative CSI-Forgery

In this subsection, we will analyze the vulnerability of SecCSI when MUs collude for CSI-forgery and propose a simple countermeasure to address this threat.

As presented in Proposition 2, both Eqs. (10) and (15) hold when $\mathbf{Q}_s = \alpha\mathbf{I}$. This condition encompasses both $\mathbf{Q}_s = \mathbf{I}$ (i.e., $\alpha = 1$, indicating the absence of CSI-forgery) and $\mathbf{Q}_s = \alpha\mathbf{I}$ with $\alpha \neq 1$. In the latter case, the MUs collaboratively report falsified CSI to produce $\mathbf{Q}_s = \alpha\mathbf{I}$, thereby circumventing the detection mechanisms presented in Eqs. (10) and (15). In what follows, we will elaborate on this collaborative CSI-forgery mode.

According to Eq. (13), if $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Q}_3 = \mathbf{Q}_4 = \mathbf{Q}_s = \alpha\mathbf{I}$, then $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_s = \alpha\mathbf{I}$ follows. In this scenario, when Eq. (10) is utilized for CSI-forgery detection, it can be expressed as:

$$[(\mathbf{P}_s + \mathbf{P}_s) - (\mathbf{P}_s + \mathbf{P}_s)]\mathbf{H}_R + (\mathbf{P}_s - \mathbf{P}_s)\mathbf{H}_D(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1} - (\mathbf{P}_s - \mathbf{P}_s)\mathbf{H}_D(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1} = \mathbf{0}. \tag{16}$$

Eq. (16) indicates that Eq. (10) is ineffective for the collaborative CSI-forgery characterized by $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_s = \alpha\mathbf{I}$. Similarly, when Eq. (15) is employed for further detection, we can derive:

$$\mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s - \mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR} = \frac{1}{\alpha}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\alpha - \mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR} = \mathbf{0}. \tag{17}$$

[6]These reliable CSIs can be directly generated at the AP, mimicking trustworthy users' CSI feedback, without any noticeable cost.

Eq. (17) indicates that Eq. (15) is also ineffective for the collaborative CSI-forgery characterized by $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Q}_3 = \mathbf{Q}_4 = \mathbf{Q}_s = \alpha\mathbf{I}$ which is equivalent to $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_s = \alpha\mathbf{I}$, either.

Next, we will present the measure that MUs cooperate to construct $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Q}_3 = \mathbf{Q}_4 = \mathbf{Q}_s = \alpha\mathbf{I}$ or equivalently $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{P}_s = \alpha\mathbf{I}$. According to Eqs. (7) and (13), the combined CSI feedback at the AP can be expressed by Eq. (18), as shown at the bottom of the page. Here, $\mathbf{H}_{Lli}$ ($l \in \{1, 2, \cdots, K_L\}$) and $\mathbf{H}_{Mmi}$ ($m \in \{1, 2, \cdots, K_M\}$) denote the estimated CSI at the l-th LU and the m-th MU during the i-th round of CEF, respectively.

From Eq. (18), we conclude that for the aforementioned collaborative CSI-forgery to occur, all users must scale their estimated CSI by the same coefficient α and report $\alpha\mathbf{H}_{Lli}$ and $\alpha\mathbf{H}_{Mmi}$ to the AP. This implies that all participating users must be colluding and coordinate their falsified CSI reports based on an identical scaling factor α, which constitutes complete collaborative CSI-forgery. Moreover, Eq. (18) suggests that as long as there is at least one LU in the system that reports genuine CSI, complete collaborative forgery cannot be achieved; then, SecCSI can effectively detect partial collaborative CSI-forgery.

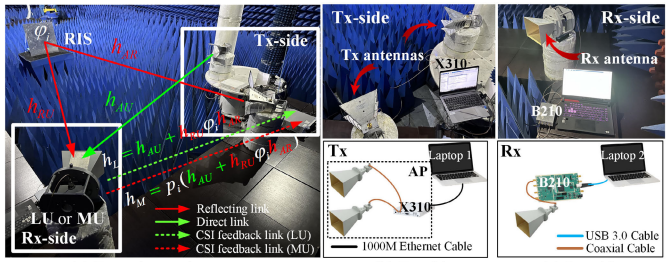
It is important to note that in practice, this requirement for the aforementioned complete collaborative CSI-forgery is extremely challenging–if not impossible–due to two main reasons: 1) The MUs cannot physically isolate LUs from the system, which allows LUs to participate in the CEF and report genuine CSI to the AP, thus violating Eq. (18); and 2) the objectives of different MUs may vary, making it unlikely for all of them to agree on a common coefficient α and establish effective collaboration. Therefore, the probability of all MUs cooperating to bypass SecCSI is very low.

To mitigate the risk of complete collaborative CSI-forgery, one can simply employ a trusted user to report genuine CSI during the CEF. This genuine CSI can either be obtained from an LU, or, more efficiently, generated by the AP to emulate an LU's CSI feedback without incurring any significant overhead (as discussed in previous subsection).

D. SecCSI's Complexity and Operational Overhead

We quantify complexity of SecCSI by the number of complex multiplications [30]. The detection process of SecCSI involves computations as outlined in Eqs. (10) and (15). According to the design of SecCSI we set $N_T = K = (K_L + K_M)N_R$ to ensure the existence of $(\mathbf{F}_3 - \mathbf{F}_4)^{-1}$, $(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1}$ and $(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1}$. For simplicity, we refer to N_T, $(K_L + K_M)N_R$, and K as N in the following analysis. The number of complex multiplications required for Eqs. (10) and (15) are $2N^2 + 4N^3$ and $N + 2N^2 + 4N^3$, respectively. Recall that Eq. (10) are used to detect forgery modes 1 to 14, while Eq. (15) is only needed for detecting forgery mode 15. Consequently, assuming that all 15 forgery modes occur with equal

$$\mathbf{F}_i = \mathbf{H}_i\mathbf{Q}_s = [\mathbf{H}_{L1i}^T \cdots \mathbf{H}_{M1i}^T \cdots \mathbf{H}_{LK_Li}^T \cdots \mathbf{H}_{MK_Mi}^T]^T \cdot \alpha\mathbf{I} = [\alpha\mathbf{H}_{L1i}^T \cdots \alpha\mathbf{H}_{M1i}^T \cdots \alpha\mathbf{H}_{LK_Li}^T \cdots \alpha\mathbf{H}_{MK_Mi}^T]^T. \tag{18}$$



(a) Experimental setup (b) Tx-side (c) Rx-side

Fig. 4. Experimental setup of SecCSI.

probability, the average computational complexity of SecCSI is $(2N^2+4N^3)+\frac{1}{15}(N+2N^2+4N^3)=\frac{1}{15}N+\frac{32}{15}N^2+\frac{64}{15}N^3$.

The operational overhead introduced by the interaction between the AP and the RIS involves two parts: 1) signaling for controlling the RIS during the estimation of $\mathbf{H}_{AR}$, and 2) signaling for controlling the RIS reflection coefficients during the four rounds of CEF. In practice, 2 bits are required to control the RIS to switch between the ON/OFF states, while the amplitude and phase shift of each RIS element are quantized with b-bit resolution [31], leading to $2Kb$ bits for coefficient configuration. During the estimation of $\mathbf{H}_{AR}$, the AP first activates the RIS and sets it to full-reflecting mode with $\phi_k = 1$ for all $k \in \{1, \cdots, K\}$. It then broadcasts a pilot signal, which reaches the RIS and is reflected back to the AP without any amplitude or phase modulation. Subsequently, the AP estimates $\mathbf{H}_{AR}$ based on this feedback pilot signal. After estimating $\mathbf{H}_{AR}$, the AP deactivates the RIS [32]. This process incurs $2+2Kb$ bits operational overhead. During the four rounds of CEF, the AP requires $2Kb$ bits each to configure $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_3$, while in the second and fourth rounds, only 1 bit per round is needed, as $\mathbf{\Phi}_2 = -\mathbf{\Phi}_1$ and $\mathbf{\Phi}_4 = -\mathbf{\Phi}_3$ can be achieved with an inversion bit. Moreover, the AP requires 8 bits to switch the RIS between the ON/OFF states during the four rounds of CEF. Therefore, the total operational overhead of SecCSI amounts to $10+4Kb$ bits. It is important to note that, in our design, K is not large, and the control of the RIS involves only switching between ON/OFF states and configuring the reflecting matrix during the four rounds of CEF. Therefore, this operational overhead is negligible compared to the volume of transmitted data.

VI. Evaluation

In this section, we first utilize the universal software radio peripheral (USRP) platform to implement SecCSI and demonstrate its validity, and then use MATLAB simulation to evaluate the detection and identification performance of SecCSI.

A. Hardware Experiment

We utilize the USRP platform to experimentally demonstrate that, with the assistance of RIS, we can establish a secure wireless environment for CSI-forgery detection, without affecting the ability of the AP to acquire the genuine CSI.

Fig. 4 shows our experimental setup. To mitigate the influence of multipath and interference on the experimental results, we conduct the experiment in an anechoic chamber.[7]

As illustrated in Fig. 4(a), our prototype system includes a Tx and a Rx representing the AP and mobile user, respectively. We use a stainless steel plate to imitate the RIS. By adjusting the angle of RIS with respect to (w.r.t.) the Tx and Rx, it can introduce a required phase shift to its incident signal As Figs. 4(b) and 4(c) show, the Tx-side consists of one USRP X310 equipped with two horn antennas, denoted as AntTx1 and AntTx2, with their main lobes targeting the RIS and the Rx, respectively. The Tx is connected to laptop 1. The Rx is implemented using a USRP B210 with a single horn antenna (denoted as AntRx1), connected to laptop 2. As Fig. 4(a) plots, AntTx1 transmits towards AntRx1, establishing a direct link between the AP and the Rx. On the other hand, AntTx2 transmits to the RIS, which then reflects the incident signal towards AntRx1, creating a reflecting link from the Tx, through the RIS, and to the Rx. We adjust RIS's angle relative to the Tx and Rx via mechanical rotation to introduce the required phase shift to the incident pilot signal [33], [34]. The main parameters used in the experiment are shown in Table II.

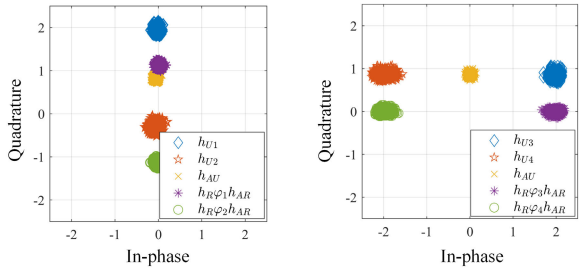
According to the design of SecCSI, the estimated CSI involving the RIS reflection in the ith round of CEF can be expressed as $\mathbf{H}_{Ui} = \mathbf{H}_{AU} + \mathbf{H}_{RU}\mathbf{\Phi}_i\mathbf{H}_{AR}$ where the subscript U of $\mathbf{H}_{Ui}$, $\mathbf{H}_{AU}$, and $\mathbf{H}_{RU}$ can be either L or M, representing LU or MU. Since we set $N_T = N_R = K = 1$ in our experiment, the estimated CSI becomes $h_{Ui} = h_{AU} + h_{RU}\varphi_i h_{AR}$, where all the variables are scalars. We configure the Rx to emulate both the MU and LU. When the Rx acts as the LU, it reports its estimated CSI $h_{Li} = h_{AU} + h_{RU}\varphi_i h_{AR} = h_{Ui}$ to the AP. In contrast, when the Rx serves as the MU, it feeds back the falsified CSI $\hat{h}_{Mi} = p_i(h_{AU} + h_{RU}\varphi_i h_{AR}) = p_i h_{Ui}$ (when $N_T = N_R = 1$, $\mathbf{P}_i$ becomes p_i). We ensure that the channel environment remains static except for the variation of φ_i during the four CEF rounds.

The experimental process is described as follows. First, we have laptop 1 control AntTx1 to transmit a pilot signal to the Rx while turning off RIS. Then, we can estimate h_{AU}, which remains constant during the four CEF rounds. Subsequently, we shut down the direct link and activate the RIS to establish a reflecting link between the Tx (transmitting with AntTx2) and Rx. We randomly place the RIS and adjust its angle w.r.t. the Tx and Rx to achieve an estimated $Ang(h_{RU}\varphi_1 h_{AR}) = 0$ where $Ang(\cdot)$ denotes the phase angle of a complex number, at the Rx. This indicates that the RIS can introduce a phase shift of θ_1 to its incident signal, where θ_1 represents the phase coefficient of the RIS in the 1st round of CEF. Then, we turn on the direct link and estimate the CSI involving the RIS reflection to obtain h_{U1}. In the 2nd round of CEF, we shut down the direct link again and adjust the angle of the RIS, yielding an estimated $Ang(h_{RU}\varphi_2 h_{AR}) = \pi$ at the Rx. This demonstrates that the RIS introduces a phase shift of $\theta_1 + \pi$ to its incoming signal. Subsequently, we turn on the direct link and estimate the CSI involving the RIS reflection to obtain h_{U2}. For the 3rd and 4th rounds of CEF, we configure the RIS's angle to yield

[7]This experimental setup does not limit the applicability of our method. SecCSI can be applied in more practical wireless environments.

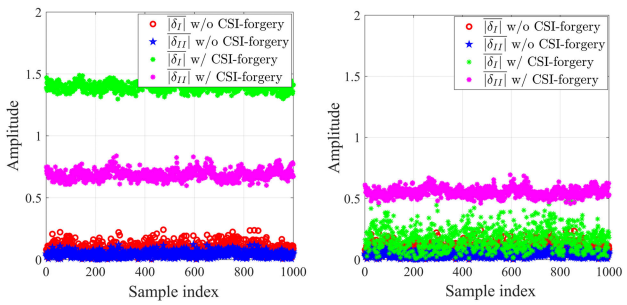
TABLE II
PARAMETER SETTINGS OF THE EXPERIMENT

Parameter	Carrier freq.	Symbol rate	Interpolation factor	Sampling rate (baseband)	Roll-off factor of raised cosine filter	Transmit gain
Value	2GHz	0.2MBaud	2	0.4MHz	0.5	15dB



(a) The 1st and 2nd rounds of CEF. (b) The 3rd and 4th rounds of CEF.

Fig. 5. Estimated channel coefficients.



(a) CSI-forgery modes 1-14 (b) CSI-forgery mode 15

Fig. 6. The average amplitude of δ_I and δ_{II} with and without CSI-forgery.

an estimated $Ang(h_{RU}\varphi_3 h_{AR}) = \frac{\pi}{2}$ at the Rx with the direct link shut down. Then, we turn on the direct link to obtain h_{U3}. Subsequently, we turn off the direct link and adjust the angle of RIS to achieve $Ang(h_{RU}\varphi_4 h_{AR}) = \frac{3\pi}{2}$, indicating that the RIS introduces a phase shift of $\theta_3 + \pi$ to its incident signal, where θ_3 represents the phase coefficient of the RIS in the 3rd round of CEF. Consequently, by activating the direct link, we can obtain h_{U4}. According to the experiment design described above, it can be ensured that $\theta_2 = \theta_1 + \pi$, $\theta_4 = \theta_3 + \pi$, and $\theta_1 \neq \theta_3$.

Fig. 5 depicts the estimated channel coefficients in a complex plane. A total of 10^3 estimated samples are plotted. As can be observed from subfigures (a) and (b), $(h_{U1} - h_{U2})/2$ and $(h_{U3} - h_{U4})/2$ can overlap with the estimated h_{AU}, which is consistent with the design of SecCSI presented in Section V. This observation also implies that the genuine h_{AU} can be concealed within the estimated h_{Ui}, thereby effectively preventing h_{AU} from being intercepted during its feedback to the AP. We can also see that the estimated $h_{RU}\varphi_1 h_{AR}$ and $h_{RU}\varphi_2 h_{AR}$ are in opposite phases, as are the $h_{RU}\varphi_3 h_{AR}$ and $h_{RU}\varphi_4 h_{AR}$. Furthermore, by comparing the two subfigures, we can verify that $\theta_3 = \theta_1 + \frac{\pi}{2}$ holds. These are in line with our experimental design.

In what follows, we will utilize the estimated CSI during the four rounds of CEF to validate that SecCSI can effectively detect potential CSI-forgery. We define $\mathbf{\Delta}_I = (\mathbf{F}_1 - \mathbf{F}_2)(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1} - (\mathbf{F}_3 - \mathbf{F}_4)(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1}$ and $\mathbf{\Delta}_{II} = (\mathbf{F}_3 - \mathbf{F}_4)^{-1}(\mathbf{F}_1 - \mathbf{F}_2) - \mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}$, which are derived from Eqs. (10) and (15), respectively, serving as indicators of CSI-forgery. Theoretically, in the absence of CSI-forgery, both $\mathbf{\Delta}_I$ and $\mathbf{\Delta}_{II}$ should be zero matrices. However, when CSI-forgery mode indexed with $\mu \in \{1, \cdots, 14\}$ occurs, $\mathbf{\Delta}_I$ becomes a non-zero matrix. In the case of CSI-forgery mode 15, $\mathbf{\Delta}_I$ is a zero matrix while $\mathbf{\Delta}_{II}$ is a non-zero matrix. Recall that in our experiment, we set $N_T = N_R = K = 1$, so the matrices mentioned above, $\mathbf{\Delta}_I$, $\mathbf{\Delta}_{II}$, $\mathbf{\Phi}_1$, $\mathbf{\Phi}_3$, and $\mathbf{H}_{AR}$ become scalars δ_I, δ_{II}, φ_1, φ_3, and h_{AR}. Then, we can have $\delta_I = \frac{f_1 - f_2}{\varphi_1 h_{AR}} - \frac{f_3 - f_4}{\varphi_3 h_{AR}}$ and $\delta_{II} = \frac{f_3 - f_4}{f_1 - f_2} - \frac{\varphi_1}{\varphi_3}$. In our experiment, we have estimated $h_{RU}\varphi_i h_{AR}$ where $i \in \{1, 2, 3, 4\}$. Since h_{RU} is immune to CSI-forgery and remains constant during the four CEF rounds, we can treat it as a static coefficient in the experiment. To avoid experimental estimation of h_{AR}, we set $h_{AR} = 1$ and substitute $\varphi_1 h_{AR}$ and $\varphi_3 h_{AR}$ in the expression of δ_I with $h_{RU}\varphi_1 h_{AR}$ and $h_{RU}\varphi_3 h_{AR}$,[8] respectively. Then, δ_I can be easily computed. Regarding δ_{II}, as both φ_1 and φ_3 are known to the AP, it can be readily obtained. We have conducted the experiment 10^3 times to collect 10^3 h_{Ui}s. Based on these h_{Ui}s, we can obtain 10^3 δ_Is and δ_{II}s without CSI-forgery. Regarding the CSI-forgery behavior, we model p_i (i.e., $\mathbf{P}_i$ when $N_T = N_R = 1$) as complex Gaussian variable with zero-mean and unit-variance. Consequently, we can apply various combinations of p_i (including setting p_i to a non-zero constant during the four rounds of CEF) to h_{Ui} to simulate the 15 CSI-forgery modes as mentioned in Section V. We can then calculate δ_I and δ_{II} with CSI-forgery accordingly.

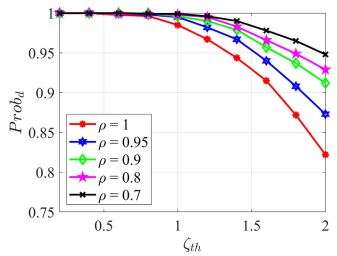
Fig. 6 plots the averaged amplitude of δ_I and δ_{II} with and without CSI-forgery over 10^3 samples, denoted as $|\overline{\delta_I}|$ w/ and w/o CSI-forgery, $|\overline{\delta_{II}}|$ w/ and w/o CSI-forgery, respectively. As the figure shows, both $|\overline{\delta_I}|$ w/o CSI-forgery and $|\overline{\delta_{II}}|$ w/o CSI-forgery are close to 0. However, due to the imperfections of CSI estimation in the experiment, they do not appear exactly 0. When the MU falsifies CSI following the CSI-forgery modes 1–14, both $|\overline{\delta_I}|$ and $|\overline{\delta_{II}}|$ obviously exceed 0. However, when the MU forges CSI following mode 15, $|\overline{\delta_I}|$ is approximately 0, meaning that Eq. (10) is incapable of identifying the presence of CSI-forgery mode 15. In this case, we need to employ Eq. (15), i.e., δ_{II}, as the criterion/indicator for further detection. As depicted in subfigure (b), when the MU employs CSI-forgery mode 15, $|\overline{\delta_{II}}|$ is obviously larger than 0, indicating the occurrence of CSI-forgery.

In summary, our USRP experiments have shown that we can utilize RIS to create a secure wireless environment, effectively preventing CSI eavesdropping and detecting potential attempts to falsify CSI.

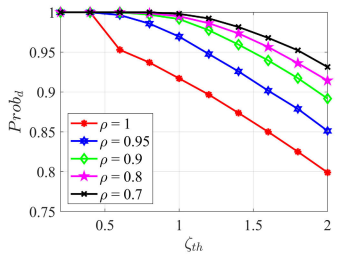
B. MATLAB Simulation

We now evaluate the performance of SecCSI using MATLAB simulation.

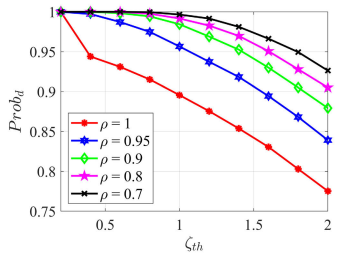
[8]This simplification may impact the detection performance of SecCSI. However, as any CSI-forgery will result in δ_I deviating from 0, the substitution only affects the magnitude of this deviation.



(a) $N_T = N_R = 4$.

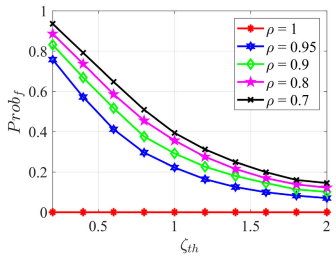


(b) $N_T = N_R = 32$.

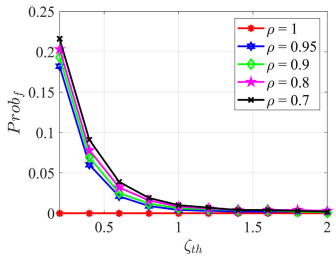


(c) $N_T = N_R = 128$.

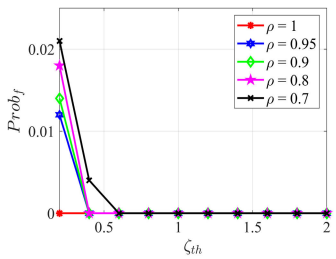
Fig. 7. Variation of $Prob_d$ with ζ_{th} under various ρs and different numbers of antennas.



(a) $N_T = N_R = 4$.



(b) $N_T = N_R = 32$.



(c) $N_T = N_R = 128$.

Fig. 8. Variation of $Prob_f$ with ζ_{th} under various ρs and different numbers of antennas.

First, we study the detection probability and false alarm probability[9] of SecCSI by setting $N_T = N_R = K \in \{4, 32, 128\}$ and $K_L + K_M = 1$. Additionally, we set $\eta_{k1} = \eta_{k3} \in \{0.1, 0.5, 1\}$ to examine the impact of the RIS absorption coefficients on detection and false alarm probabilities under the setting $N_T = N_R = K = 4$. We model $\mathbf{P}_i$ ($i \in \{1, 2, 3, 4\}$) as complex Gaussian matrix with zero-mean and unit-variance to characterize CSI-forgery behavior. This encompasses both known and unknown CSI-forgery attacks, making the generation of $\mathbf{P}_i$ according to specific CSI-forgery attacks unnecessary. We study the scenario without CSI-forgery and a total of 15 CSI-forgery behaviors, as discussed in Section V. Second, we set $N_T = K = 4$, $N_R = 1$, $K_L = 3$, and $K_M = 1$ to evaluate the identification accuracy of MU in a multi-user system using SecCSI. In this configuration, $\mathbf{P}_i = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_{3\times 1} \\ [w_{1i}\ w_{2i}\ w_{3i}] & w_{4i} \end{bmatrix}$, where $\mathbf{I}_3$ denotes the 3×3 identity matrix and $\mathbf{0}_{3\times 1}$ represents the 3×1 zero column vector, respectively. w_{ji} ($j \in \{1, 2, 3, 4\}$) is modeled as a complex Gaussian random variable with zero mean and unit variance. Finally, we evaluate the detection performance of SecCSI against three representative CSI-forgery attacks discussed in Section IV (i.e., Sniff attack, USE attack and RIS Jamming attack), along with the Random attack (where we model w_{ji} as the complex Gaussian random variable with zero mean and unit variance).

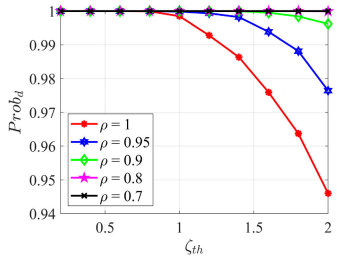
In practical applications, the imperfectness of channel estimation can impact the detection performance of SecCSI. We use Eq. (19) to characterize the imperfections in channel estimation as [35]:

$$\tilde{\mathbf{H}} = \rho \mathbf{H} + \sqrt{1 - \rho^2}\mathbf{E}, \tag{19}$$

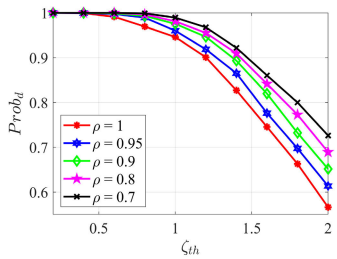
where $\mathbf{H}$ and $\tilde{\mathbf{H}}$ denote the perfect CSI and imperfect CSI, respectively. $\rho \in (0, 1]$ represents the degree of imperfection in channel estimation. Specifically, when $\rho = 1$, $\tilde{\mathbf{H}} = \mathbf{H}$ holds, indicating accurate channel estimation. $\mathbf{E}$ is a $N_R \times N_T$ complex Gaussian matrix with zero-mean and unit-variance complex Gaussian variables as its elements.

1) Detection Probability and False Alarm Probability in a Single User System: In the simulation of SecCSI's detection performance, we generate 5×10^4 sets of $\mathbf{H}_R$, $\mathbf{H}_{AR}$, and $\mathbf{H}_D$ independently according to the System Model Section. Consequently, the ideal estimated CSI can be calculated as $\mathbf{H}_{Ui} = \mathbf{H}_{AU} + \mathbf{H}_{RU}\mathbf{\Phi}_i\mathbf{H}_{AR}$ where $i \in \{1, 2, 3, 4\}$ and the subscript U can be either L (without CSI-forgery) or M (with CSI-forgery). As for CSI-forgery, we apply $\mathbf{P}_i$ to the estimated $\mathbf{H}_{Mi}$ to yield $\hat{\mathbf{H}}_{Mi}$. Since there are 15 CSI-forgery modes in total, for each set of $\mathbf{H}_R$, $\mathbf{H}_{AR}$, and $\mathbf{H}_D$, we will obtain 15 CSI-forgery samples in total. Recall that $\phi_{ki} = \eta_{ki}e^{j\theta_{ki}}$, where $k \in \{1, \cdots, K\}$, denotes the kth reflecting coefficient of $\mathbf{\Phi}_i$ in the ith round of CEF. We randomly select η_{k1} and

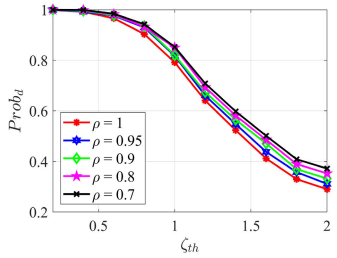
[9]Since the probability of missed detection can be obtained by subtracting the detection probability from 1, we omit its separate discussion.



(a) $\eta_{k1} = \eta_{k3} = 0.1$.

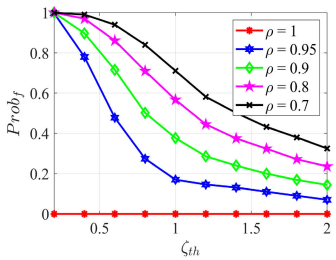


(b) $\eta_{k1} = \eta_{k3} = 0.5$.

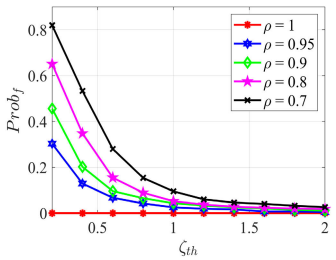


(c) $\eta_{k1} = \eta_{k3} = 1$.

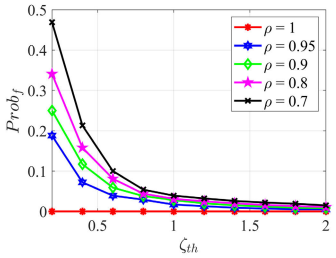
Fig. 9. Variation of $Prob_d$ with ζ_{th} under various ρs and different RIS's absorption coefficients.



(a) $\eta_{k1} = \eta_{k3} = 0.1$.



(b) $\eta_{k1} = \eta_{k3} = 0.5$.



(c) $\eta_{k1} = \eta_{k3} = 1$.

Fig. 10. Variation of $Prob_f$ with ζ_{th} under various ρs and different RIS's absorption coefficients.

η_{k3} from the interval $[0, 1)$, and θ_{k1} and θ_{k3} from the range $[0, \pi]$, satisfying $\mathbf{\Phi}_1 = -\mathbf{\Phi}_2$, $\mathbf{\Phi}_3 = -\mathbf{\Phi}_4$, and $\mathbf{\Phi}_1 \neq \mathbf{\Phi}_3$. As a result, we can compute 5×10^4 pairs of $\mathbf{\Delta}_I$ and $\mathbf{\Delta}_{II}$ for the case without CSI-forgery, and for each CSI-forgery mode. We then use $\zeta_I = \|\mathbf{\Delta}_I\|_F/(N_R N_T)$ and $\zeta_{II} = \|\mathbf{\Delta}_{II}\|_F/(N_R N_T)$, where $\|\cdot\|_F$ represents the Frobenius norm, as the indicator of CSI-forgery. We employ ζ_{th} as the threshold for detecting CSI-forgery.[10] Specifically, if $\zeta_I > \zeta_{th}$, we can conclude that CSI-forgery mode indexed with $\mu \in \{1, \cdots, 14\}$ has occurred. Alternatively, if $\zeta_I < \zeta_{th}$ but $\zeta_{II} > \zeta_{th}$, we can determine that CSI-forgery mode 15 has occurred. It is important to note that the values of ζ_I and ζ_{II} are primarily influenced by three factors: 1) The disparities among the CSI-forgery matrices $\mathbf{P}_i (i \in \{1, 2, 3, 4\})$; 2) The amplification of channel estimation errors due to increased matrix condition numbers, which exacerbates inaccuracies in matrix inversion calculations; and 3) The diminishing effect on both ζ_I and ζ_{II} as N_R and N_T increase.

Fig. 7 plots the variation of the detection probability, denoted as $Prob_d$, of SecCSI with ζ_{th} under various ρs and different numbers of antennas, where only the samples with CSI-forgery are studied. As the figure shows, the curves of $Prob_d$ decrease with the increase of ζ_{th}. This is because the deviation of ζ_I and ζ_{II} from zero is influenced by both the CSI-forgery behavior and CSI estimation accuracy. Therefore, as ζ_{th} grows large, even in the presence of CSI-forgery, SecCSI may not be able to detect its occurrence, decreasing $Prob_d$. Moreover, with a fixed ζ_{th}, $Prob_d$ grows as ρ rises. This is because, as ρ increases, the imperfections in CSI estimation grow, leading to an increase in the deviation of ζ_I and ζ_{II} from zero. Consequently, even in the absence of CSI-forgery, SecCSI may incorrectly identify the presence of CSI-forgery. However, it is worth noting that even with consideration of imperfect CSI estimation, SecCSI can exhibit excellent detection performance. Specifically, when $N_T = N_R = 4$, by setting $\zeta_{th} \leq 1.3$, $Prob_d$ can reach over 95% with SecCSI even when using poor channel estimation methods.

Furthermore, we can observe from Fig. 7 that $Prob_d$ slightly decreases with increasing N_T and N_R. This is because growing condition number amplifies the impact of CSI imperfections, further increasing both ζ_I and ζ_{II}. Moreover, the normalization by $N_R N_T$ in the calculations of ζ_I and ζ_{II} causes their values to diminish as N_R and N_T increase. As a result, the decrease in $Prob_d$ with the rise in N_T and N_R remains limited. Additionally, we observe from Fig. 7 that under perfect CSI (i.e., $\rho = 1$), the influence of CSI imperfections on matrix inversion calculations is eliminated. Consequently,

[10] According to the definitions of ζ_I and ζ_{II}, the selection of ζ_{th} depends on the channel conditions, the number of antennas at both the AP and users, as well as the configuration of the RIS.

$Prob_d$ decreases significantly when ζ_{th} is small as N_R and N_T increase.

Fig. 8 illustrates the variation of false alarm probability, denoted as $Prob_f$, of SecCSI with ζ_{th} under various ρs and different numbers of antennas, where only the samples without CSI-forgery are considered. As the figure plots, given $\rho = 1$, $Prob_f$ remains constant at 0 irrespective of the increase in ζ_{th}. This implies that as long as the channel estimation is accurate, SecCSI will not mistake the absence of CSI-forgery for the presence of CSI-forgery. Regarding the other curves of $Prob_f$, they gradually decrease as ζ_{th} increases. This is because, when there is no CSI-forgery, only the inaccuracy of CSI estimation leads to the deviation of ζ_I and ζ_{II} from 0. Therefore, as ζ_{th} enlarges, it becomes less likely to mistake the channel estimation error for CSI-forgery. Additionally, with a fixed ζ_{th}, $Prob_f$ grows as ρ reduces, indicating that as imperfections in CSI estimation increase, SecCSI may be more inclined to mistake the absence of CSI-forgery for the existence of CSI-forgery. However, it is worth noting that as long as we set $\zeta_{th} > 1.4$, $Prob_f$ under $N_T = N_R = 4$ can be less than 25% even when the accuracy of channel estimation is poor. In practice, we can utilize accurate CSI estimation methods to ensure an acceptable $Prob_f$.

Additionally, we can observe from Fig. 8 that $Prob_f$ significantly decreases as N_T and N_R increase. This occurs because $Prob_f$ is evaluated in the absence of CSI-forgery, where $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_4 = \mathbf{I}$, thus eliminating the influence of disparities among the CSI-forgery matrices on ζ_I and ζ_{II}. In this scenario, although the increase in N_R and N_T results in high condition numbers that amplify the inaccuracies of matrix inversion calculations — thereby increasing both ζ_I and ζ_{II} — this amplification diminishes as N_R and N_T become very large due to the normalization of ζ_I and ζ_{II} by $N_R N_T$.

Next, we examine the impact of RIS's coefficients on $Prob_d$ and $Prob_f$ under $N_T = N_R = K = 4$, $K_M = 1$, and $\eta_{k1} = \eta_{k3} \in \{0.1, 0.5, 1\}$ where $k \in \{1, 2, \cdots, K\}$.

Fig. 9 illustrates the variation of $Prob_d$ with ζ_{th} under $\eta_{k1} = \eta_{k3} \in \{0.1, 0.5, 1\}$. As the figure shows, $Prob_d$ decreases as the RIS absorption coefficients increase. This occurs because smaller values of η_{k1} and η_{k3} lead to higher gains in the inverse diagonal reflecting matrices $\mathbf{\Phi}_1^{-1}$ and $\mathbf{\Phi}_3^{-1}$. Consequently, this increases both $\|\mathbf{H}_D(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1}\|_F$ and $\|\mathbf{H}_D(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1}\|_F$, which in turn enlarges of ζ_I, ζ_{II}, and $Prob_d$.

Fig. 10 plots the variation of $Prob_f$ with ζ_{th} under $\eta_{k1} = \eta_{k3} \in \{0.1, 0.5, 1\}$. As the figure shows, $Prob_f$ decreases with increasing η_{k1} and η_{k3}. This occurs because smaller values of η_{k1} and η_{k3} lead to higher $\mathbf{\Phi}_1^{-1}$ and $\mathbf{\Phi}_3^{-1}$. Consequently, this amplifies the channel estimation error, which in turn enlarges ζ_I, ζ_{II}, and $Prob_f$.

The results presented in Figs. 9 and 10 suggest that while a higher absorption coefficient at the RIS enhances SecCSI's ability to detect CSI-forgery behaviors, it also increases the likelihood of misclassifying a LU as malicious in the absence of CSI-forgery attacks. Therefore, the selection of the RIS absorption coefficient is essential to balance $Prob_d$ and $Prob_f$.

2) Identification Probability in a Multi-User System: We now evaluate the MU identification performance of SecCSI under the influence of CSI imperfections. As mentioned before, we consider a system consisting of one MU and three LUs. The MU performs each of the 15 CSI-forgery modes 5×10^4 times w.r.t. the 5×10^4 estimated $\mathbf{h}_{Mi}$ where $i \in \{1, 2, 3, 4\}$. To exclusively evaluate the identification performance of SecCSI, we assume that the AP has correctly detected the presence of CSI-forgery. Consequently, as long as the AP can use the method presented in Section V to correctly identify that the MU provides falsified CSI, we increment the count of successful MU identification by 1.

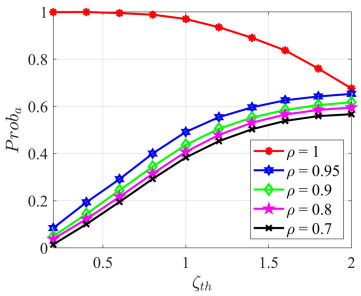


Fig. 11. Variation of $Prob_a$ with ζ_{th} under various ρs.

Fig. 11 plots the variation of the accuracy of MU identification, defined as the probability that the AP can accurately attribute CSI-forgery to the MU, denoted as $Prob_a$, of SecCSI with ζ_{th} under various ρs. As the figure shows, when $\rho = 1$, $Prob_a$ decreases as ζ_{th} increases, whereas the remaining $Prob_a$ curves for $\rho < 1$ rise as ζ_{th} grows. This is because $Prob_a$ is proportional to the product of $Prob_d$ and $1 - Prob_f$. Then, as illustrated in Fig. 8(a), $Prob_f$ remains constant at 0 regardless of the changes in ζ_{th} when $\rho = 1$, letting $Prob_a$ rely solely on $Prob_d$. Therefore, $Prob_a$ exhibits the same pattern as $Prob_d$ depicted in Fig. 7(a) when $\rho = 1$. Nevertheless, when $\rho < 1$, $Prob_a$ is affected by both $Prob_d$ and $Prob_f$. In such a case, since the absolute value of the slope of $1 - Prob_f$ is greater than that of $Prob_d$, as can be deduced from the comparison of Figs. 7(a) and 8(a), $Prob_a$ tends to increase with an increase in ζ_{th}. Furthermore, with a fixed ζ_{th}, $Prob_a$ decreases as ρ decreases, indicating that as the imperfection in CSI estimation rises, the accuracy of SecCSI in identifying the MU deteriorates. This is because a higher ρ can result in a greater likelihood of LU being misclassified as MU.

3) Detection Probability Under Different CSI-Forgery Attacks in a Multi-User System: Finally, we simulate the detection probability of SecCSI concerning Sniff attack, USE attack and RIS Jamming attack (all of which can be classified into the CSI-forgery modes indexed with 1), along with Random attack, under perfect CSI (i.e., $\rho = 1$). We consider a system consisting of one MU and three LUs. The MU performs each of the aforementioned CSI-forgery attacks 5×10^4 times based on 5×10^4 generations of $\mathbf{h}_M$.

As Fig. 12 shows, SecCSI achieves the highest $Prob_d$ for the Sniff attack, followed by the Random attack, then the RIS Jamming attack, while $Prob_d$ for the USE attack is the lowest. This variation arises because different types of CSI-forgery attacks have distinct $\mathbf{P}_i$ ($i \in \{1, 2, 3, 4\}$), leading to different values of ζ_I and ζ_{II} (details can be found in the third paragraph on page 11), as well as $Prob_d$. Since the Random attack models w_{ji} ($j \in \{1, 2, 3, 4\}$) as a complex Gaussian random

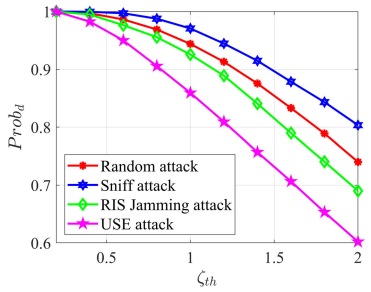


Fig. 12. Variation of $Prob_d$ with ζ_{th} under different CSI-forgery attacks.

variable, its $Prob_d$ is averaged across all possible CSI-forgery behaviors.

Based on the simulation results and the fact that CSI estimation is usually imperfect, we recommend opting for a larger ζ_{th} to ensure lower $Prob_f$ and higher $Prob_a$, whereas a smaller ζ_{th} is preferable to achieve a higher $Prob_d$. So, we can first employ a smaller ζ_{th} to ensure any potential CSI-forgery can be detected, and then switch to a larger ζ_{th} in the MU identification phase to ensure the attacker can be correctly identified with a high probability. Above all, a high-accuracy CSI estimation method is essential for ensuring the effectiveness of SecCSI.

It is important to note that conducting a quantitative comparison between SecCSI and other existing methods is challenging. This is because: 1) existing countermeasures are based on various principles, leading to differences in hardware requirements, resource consumption, and other factors, which complicates their comparisons under common conditions; 2) some countermeasures are designed for specific system settings, making their applicability to more general settings uncertain thus preventing comparisons within a common system framework; and 3) some literature primarily focuses on the design of attacks, with limited discussion of countermeasures, resulting in difficulties in conducting a thorough performance evaluation of these methods. However, through a qualitative comparison with the methods presented in Table I, we can conclude that SecCSI is the only countermeasure that comprehensively prevents CSI eavesdropping, effectively detects both known and unknown CSI-forgery attacks, and identifies the attackers without incurring additional hardware or power overhead.

VII. Conclusion

In this paper, we explored the use of RIS in establishing a secure wireless environment to defend against potential CSI-forgery from malicious attackers. We first developed a general model to subsume existing CSI-forgery attacks as special instances. Then, by using this model, we proposed SecCSI for detecting CSI-forgery and identifying the MU responsible for falsifying CSI by dynamically configuring the reflection coefficients of the RIS. We conducted USRP experiments and MATLAB simulations to demonstrate the validity and evaluate the performance of SecCSI. Our theoretical analysis, experimental and numerical evaluations have shown SecCSI to effectively prevent CSI interception, detect the CSI-forgery attacks and identify the attackers.

Appendix A
Proof of Proposition 1

According to Eq. (7), the AP receives the combined feedback CSI in the first round of CEF as $\mathbf{F}_1 = \mathbf{P}_1(\mathbf{H}_R\mathbf{\Phi}_1\mathbf{H}_{AR} + \mathbf{H}_D)$. However, since the AP lacks prior knowledge of $\mathbf{H}_R$ and $\mathbf{H}_D$, it cannot derive $\mathbf{P}_1$ from $\mathbf{F}_1$. Consequently, subsequent CSI feedback from users is necessary to detect the presence of $\mathbf{P}_i$ through comparison of the feedback, indicating that at least two rounds of CEF are required. In the following analysis, we will explore the limitations associated with using only two or three rounds of CEF.

(1) For two rounds of CEF, we can derive $\mathbf{F}_1 - \mathbf{F}_2$ from the successive two rounds of CEF as:

$$\mathbf{F}_1 - \mathbf{F}_2 = (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{H}_D + (\mathbf{P}_1\mathbf{H}_R\mathbf{\Phi}_1 - \mathbf{P}_2\mathbf{H}_R\mathbf{\Phi}_2)\mathbf{H}_{AR}. \quad (20)$$

If $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{I}$ (i.e., in the absence of CSI-forgery), we can rewrite Eq. (20) as:

$$\mathbf{F}_1 - \mathbf{F}_2 = \mathbf{H}_R(\mathbf{\Phi}_1 - \mathbf{\Phi}_2)\mathbf{H}_{AR}. \quad (21)$$

It is evident that setting $\mathbf{\Phi}_1 = \mathbf{\Phi}_2$ results in Eq. (21) being a zero matrix. Therefore, substituting $\mathbf{\Phi}_1 = \mathbf{\Phi}_2$ into Eq. (20) allows us to detect any CSI-forgery behavior characterized by $\mathbf{P}_1 \neq \mathbf{P}_2$ by examining whether $\mathbf{F}_1 - \mathbf{F}_2 = \mathbf{0}$ holds.

However, when $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_s \neq \mathbf{I}$, the previous detection scheme becomes ineffective. In this scenario, Eq. (20) becomes:

$$\mathbf{F}_1 - \mathbf{F}_2 = \mathbf{P}_s\mathbf{H}_R(\mathbf{\Phi}_1 - \mathbf{\Phi}_2)\mathbf{H}_{AR}. \quad (22)$$

Since $\mathbf{H}_R$ is unknown at the AP, determining whether $\mathbf{P}_s = \mathbf{I}$ based on Eq. (22) is impossible.

(2) For three rounds of CEF, without loss of generality, we can compare $\mathbf{F}_2$ and $\mathbf{F}_3$ with $\mathbf{F}_1$ to obtain:

$$\begin{cases}\mathbf{F}_1 - \mathbf{F}_2 = (\mathbf{P}_1 + \mathbf{P}_2)(\mathbf{H}_R\mathbf{\Phi}_1\mathbf{H}_{AR}) + (\mathbf{P}_1 - \mathbf{P}_2)\mathbf{H}_D \\ \mathbf{F}_1 - \mathbf{F}_3 = (\mathbf{P}_1\mathbf{H}_R\mathbf{\Phi}_1 - \mathbf{P}_3\mathbf{H}_R\mathbf{\Phi}_3)\mathbf{H}_{AR} + (\mathbf{P}_1 - \mathbf{P}_3)\mathbf{H}_D\end{cases} \quad (23)$$

If $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{I}$ (i.e., in the absence of CSI-forgery), We can derive Eq. (24) from Eq. (23) as:

$$\begin{aligned}&(\mathbf{F}_1 - \mathbf{F}_2)(\mathbf{\Phi}_1\mathbf{H}_{AR})^{-1} - (\mathbf{F}_1 - \mathbf{F}_3)(\mathbf{\Phi}_3\mathbf{H}_{AR})^{-1} \\ &= 2\mathbf{H}_R - \mathbf{H}_R\mathbf{\Phi}_1\mathbf{\Phi}_3^{-1} + \mathbf{H}_R = \mathbf{H}_R(3\mathbf{I} - \mathbf{\Phi}_1\mathbf{\Phi}_3^{-1}).\end{aligned} \quad (24)$$

Setting $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_3$ such that $\mathbf{\Phi}_1\mathbf{\Phi}_3^{-1} = 3\mathbf{I}$ (i.e., $\mathbf{\Phi}_1 = 3\mathbf{\Phi}_3$) shows that Eq. (24) yields a zero matrix in the absence of CSI-forgery. Conversely, under CSI-forgery modes indexed from 1 to 14, Eq. (24) remains a non-zero matrix, indicating its utility for detecting these 14 forgery modes using three rounds of CEF.

However, when $\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P}_3 = \mathbf{P}_s \neq \mathbf{I}$, we can derive Eq. (25) from Eq. (13) and Eq. (23) as:

$$\begin{cases}\mathbf{F}_1 - \mathbf{F}_2 = 2\mathbf{H}_R\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s \\ \mathbf{F}_1 - \mathbf{F}_3 = \mathbf{H}_R(\mathbf{\Phi}_1 - \mathbf{\Phi}_3)\mathbf{H}_{AR}\mathbf{Q}_s\end{cases} \quad (25)$$

Then, we can have:

$$(\mathbf{F}_1 - \mathbf{F}_3)^{-1}(\mathbf{F}_1 - \mathbf{F}_2) = \mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}(\mathbf{\Phi}_1 - \mathbf{\Phi}_3)^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s. \quad (26)$$

Substituting $\mathbf{\Phi}_1 = 3\mathbf{\Phi}_3$ into Eq. (26), we have $\mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\left(\frac{2}{3}\mathbf{\Phi}_1\right)^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s = \frac{3}{2}\mathbf{I}$. This equation holds regardless of whether $\mathbf{Q}_s = \mathbf{I}$. Consequently, SecCSI fails to

$$(\mathbf{F}_3-\mathbf{F}_4)^{-1}(\mathbf{F}_1-\mathbf{F}_2)=[\mathbf{H}_R\mathbf{\Phi}_3\mathbf{H}_{AR}(\mathbf{Q}_3+\mathbf{Q}_4)+\mathbf{H}_D(\mathbf{Q}_3-\mathbf{Q}_4)]^{-1}[\mathbf{H}_R\mathbf{\Phi}_1\mathbf{H}_{AR}(\mathbf{Q}_1+\mathbf{Q}_2)+\mathbf{H}_D(\mathbf{Q}_1-\mathbf{Q}_2)]. \quad (29)$$

distinguish between the presence and absence of CSI-forgery when $\mathbf{P}_1=\mathbf{P}_2=\mathbf{P}_3=\mathbf{P}_s\neq\mathbf{I}$ using three rounds of CEF.

Based on the above analysis, we conclude that neither two nor three rounds of CEF are sufficient for detecting all CSI-forgery modes. Therefore, Proposition 1 follows. ∎

Appendix B

According to Eq. (13), the combined feedback CSI at the AP can be represented as:

$$\mathbf{F}_i=(\mathbf{H}_R\mathbf{\Phi}_i\mathbf{H}_{AR}+\mathbf{H}_D)\,\mathbf{Q}_i. \quad (27)$$

Then, by substituting Eq. (27) into Eqs. (8) and (11), we can have:

$$\begin{cases}\mathbf{F}_1-\mathbf{F}_2=(\mathbf{H}_R\mathbf{\Phi}_1\mathbf{H}_{AR})(\mathbf{Q}_1+\mathbf{Q}_2)+\mathbf{H}_D(\mathbf{Q}_1-\mathbf{Q}_2)\\ \mathbf{F}_3-\mathbf{F}_4=(\mathbf{H}_R\mathbf{\Phi}_3\mathbf{H}_{AR})(\mathbf{Q}_3+\mathbf{Q}_4)+\mathbf{H}_D(\mathbf{Q}_3-\mathbf{Q}_4)\end{cases}. \quad (28)$$

By left-multipying $(\mathbf{F}_3-\mathbf{F}_4)^{-1}$ by $(\mathbf{F}_1-\mathbf{F}_2)$, we can obtain Eq. (29), as shown at the top of the page.

Since we assume that $\mathbf{Q}_1=\mathbf{Q}_2=\mathbf{Q}_3=\mathbf{Q}_4=\mathbf{Q}_s\neq\alpha\mathbf{I}$ where α is non-zero scalar, we can rewrite Eq. (29) as:

$$(\mathbf{F}_3-\mathbf{F}_4)^{-1}(\mathbf{F}_1-\mathbf{F}_2)=\mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s. \quad (30)$$

Substituting Eq. (30) into Eq. (15), we can obtain:

$$\mathbf{Q}_s^{-1}\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}\mathbf{Q}_s-\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}=\mathbf{0}. \quad (31)$$

Then, we define $\mathbf{A}=\mathbf{H}_{AR}^{-1}\mathbf{\Phi}_3^{-1}\mathbf{\Phi}_1\mathbf{H}_{AR}$ and can simplify Eq. (31) to:

$$\mathbf{Q}_s^{-1}\mathbf{A}\mathbf{Q}_s=\mathbf{A}. \quad (32)$$

It can be easily seen that for any non-zero matrix $\mathbf{A}$, as long as $\mathbf{A}\mathbf{Q}_s=\mathbf{Q}_s\mathbf{A}$ is satisfied, Eq. (32) can hold true. However, the necessary and sufficient condition for $\mathbf{A}\mathbf{Q}_s=\mathbf{Q}_s\mathbf{A}$ to hold is that both matrices share the same n independent eigenvectors, where n denotes the order of $\mathbf{A}$ and $\mathbf{Q}_s$ [36]. Sicne $\mathbf{A}$ is random, $\mathbf{Q}_s$ should take the forms given by Eq. (32) to satisfy the aforementioned condition [37]:

$$\mathbf{Q}_s=\begin{cases}\mathbf{B}^{-1}\mathbf{A}\mathbf{B}\\ a_0\mathbf{I}+a_1\mathbf{A}^1+\cdots+a_L\mathbf{A}^L,\end{cases} \quad (33)$$

where $\mathbf{B}$ represents an arbitrary invertible matrix and a_l ($l\in\{0,1,\cdots,L\}$ is an arbitrary coefficient.

Since accurately acquiring $\mathbf{A}$ is impossible for the MU in practice and $\mathbf{Q}_s$ is a non-zero matrix, $\mathbf{Q}_s$ can only take the form of $a_0\mathbf{I}$ to ensure that Eq. (31) holds. However, this requirement is contradictory to the assumption of $\mathbf{Q}_s\neq\alpha\mathbf{I}$. Therefore, Proposition 2 follows. ∎

REFERENCES

[1] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 763–776,Jan. 2017.

[2] N. Aldaghri and H. Mahdavifar, "Physical layer secret key generation in static environments," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2692–2705, 2020.

[3] Q. Wang, K. Xu, and K. Ren, "Cooperative secret key generation from phase estimation in narrowband fading channels," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1666–1674, Oct. 2012.

[4] Y.-C. Tung, S. Han, D. Chen, and K. G. Shin, "Vulnerability and protection of channel state information in multiuser MIMO networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Nov. 2014, pp. 775–786.

[5] X. Wang, Y. Liu, X. Lu, S. Lv, Z. Shi, and L. Sun, "On eavesdropping attacks and countermeasures for MU-MIMO systems," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2017, pp. 40–45.

[6] Y.-C. Tung, K. G. Shin, and K.-H. Kim, "Analog man-in-the-middle attack against link-based packet source identification," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jul. 2016, pp. 331–340.

[7] S. Wang, Z. Chen, Y. Xu, Q. Yan, C. Xu, and X. Wang, "On user selective eavesdropping attacks in MU-MIMO: CSI forgery and countermeasure," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Apr. 2019, pp. 1963–1971.

[8] T. Hou et al., "MUSTER: Subverting user selection in MU-MIMO networks," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2022, pp. 140–149.

[9] X. Zhou, B. Maham, and A. Hjorungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 903–907, Mar. 2012.

[10] L. Li et al., "Electromagnetic reprogrammable coding-metasurface holograms," *Nature Commun.*, vol. 8, no. 1, p. 197, Aug. 2017.

[11] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Jan. 2020.

[12] X. Cheng et al., "Joint optimization for RIS-assisted wireless communications: From physical and electromagnetic perspectives," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 606–620, Jan. 2022.

[13] T. Jiang and W. Yu, "Interference nulling using reconfigurable intelligent surface," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 5, pp. 1392–1406, May 2022.

[14] C. Pan et al., "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, Aug. 2020.

[15] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Artificial-noise-aided secure MIMO wireless communications via intelligent reflecting surface," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7851–7866, Dec. 2020.

[16] J. Luo, F. Wang, S. Wang, H. Wang, and D. Wang, "Reconfigurable intelligent surface: Reflection design against passive eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3350–3364, May 2021.

[17] P. Staat et al., "IRShield: A countermeasure against adversarial physical-layer wireless sensing," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2022, pp. 1705–1721.

[18] H. Niu, X. Lei, J. An, L. Zhang, and C. Yuen, "On the efficient design of stacked intelligent metasurfaces for secure SISO transmission," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 60–70, 2025.

[19] H. Niu, Y. Xiao, X. Lei, L. Dan, W. Xiang, and C. Yuen, "Reconfigurable intelligent surface-assisted passive beamforming attack," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 8236–8247, 2024.

[20] G. Li et al., "RIS-jamming: Breaking key consistency in channel reciprocity-based key generation," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 5090–5105, 2024.

[21] J. Yang, X. Ji, F. Wang, K. Huang, and L. Guo, "A novel pilot spoofing scheme via intelligent reflecting surface based on statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 12847–12857, Dec. 2021.

[22] K.-W. Huang and H.-M. Wang, "Intelligent reflecting surface aided pilot contamination attack and its countermeasure," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 345–359, Jan. 2021.

[23] Y. Liu, S. Zhang, F. Gao, J. Tang, and O. A. Dobre, "Cascaded channel estimation for RIS assisted mmWave MIMO transmissions," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2065–2069, Sep. 2021.

[24] G. S. Park and H. Song, "Cooperative base station caching and X2 link traffic offloading system for video streaming over SDN-enabled 5G networks," *IEEE Trans. Mobile Comput.*, vol. 18, no. 9, pp. 2005–2019, Sep. 2019.

[25] C. Song, "Leakage rate analysis for artificial noise assisted massive MIMO with non-coherent passive eavesdropper in block-fading," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2111–2124, Apr. 2019.

[26] C. Liu, J. Zhou, Y. Gao, D. Qiao, and H. Qian, "IRS-aided secure communications over an untrusted AF relay system," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 8620–8633, Dec. 2023.

[27] W. Pamungkas, T. Suryani, Wirawan, A. Affandi, R. Ananda, and J. Hendry, "An improved method to detect coherence time in wireless communications channel based on auto correlation functions," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 4, pp. 4039–4055, Apr. 2023.
[28] J. Choi, S. Choi, and K. B. Lee, "Sounding node set and sounding interval determination for IEEE 802.11ac MU-MIMO," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10069–10074, Dec. 2016.
[29] X. Ma, Q. Gao, J. Wang, V. Marojevic, and J. H. Reed, "Dynamic sounding for multi-user MIMO in wireless LANs," *IEEE Trans. Consum. Electron.*, vol. 63, no. 2, pp. 135–144, May 2017.
[30] A. Aminjavaheri, A. Farhang, A. Rezazadehreyhani, L. E. Doyle, and B. Farhang-Boroujeny, "OFDM without CP in massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7619–7633, Nov. 2017.
[31] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
[32] C. Hu, L. Dai, S. Han, and X. Wang, "Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7736–7747, Nov. 2021.
[33] H. Lu, Z. Yu, Y. Zeng, S. Ma, S. Jin, and R. Zhang, "Wireless communication with flexible reflector: Joint placement and rotation optimization for coverage enhancement," *IEEE Trans. Wireless Commun.*, vol. 24, no. 10, pp. 8252–8266, Oct. 2025.
[34] Z. Yu, C. Feng, Y. Zeng, T. Li, and S. Jin, "Wireless communication using metal reflectors: Reflection modelling and experimental verification," in *Proc. IEEE Int. Conf. Commun.*, May 2023, pp. 4701–4706.
[35] Z. Li, Y. Liu, K. G. Shin, J. Li, F. Guo, and J. Liu, "Design and adaptation of multi-interference steering," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3329–3346, Jul. 2019.
[36] G. Strang, *Introduction to Linear Algebra*. Cambridge, MA, USA: Wellesley-Cambridge, 1993.
[37] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2013.

Yicheng Liu (Graduate Student Member, IEEE) is currently working toward the Ph.D. degree with the School of Cyber Engineering, Xidian University. His research interests include wireless communication, physical layer security, and interference management.

Zhao Li (Senior Member, IEEE) is currently an Associate Professor with the School of Cyber Engineering, Xidian University. He has published more than 60 technical papers at premium international journals and conferences, such as IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE INFOCOM. His research interests include wireless communication, 5G communication systems, resource allocation, interference management, the IoT, and physical layer security.

Kang G. Shin (Life Fellow, IEEE) is the Kevin and Nancy O'Connor Professor of Computer Science and Founding Director of the Real-Time Computing Laboratory with the Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, Michigan. His current research focuses on safe and secure embedded real-time and cyber-physical systems as well as QoS-sensitive computing and networking. He has supervised the completion of 93 Ph.D. and authored/coauthored about 1,000 technical articles, a text book and about 60 patents or invention disclosures, and received, numerous awards, including 2023 IEEE TCCPS Technical Achievement Award, 2023 SIGMOBILE Test-of-Time Award, 2019 Caspar Bowden Award for Outstanding Research in Privacy Enhancing Technologies, and the Best Paper Awards from 2023 VehicleSec, 2011 ACM International Conference on Mobile Computing and Networking(MobiCom'11), the 2011 IEEE International Conference on Autonomic Computing, 2010 and 2000 USENIX Annual Technical Conferences, as well as the2003 IEEE Communications Society William R. Bennett Prize Paper Award and the 1987 Outstanding IEEE TRANSACTIONS OF AUTOMATIC CONTROL Paper Award. He has also received several institutional awards, including the Research Excellence Award, in 1989, Outstanding Achievement Award, in 1999, Distinguished Faculty Achievement Award in 2001, and Stephen Attwood Award in 2004 from The University of Michigan (the highest honor bestowed to Michigan Engineering faculty); a Distinguished Alumni Award of the College of Engineering, Seoul National University in 2002; 2003 IEEE RTC Technical Achievement Award;and 2006 Ho-Am Prize in Engineering (the highest honor bestowed to Korean origin engineers). He has chaired Michigan Computer Science and Engineering Division for four years starting 1991, and also several major conferences, including 2009 ACM MobiCom, and 2005 ACM/USENIX MobiSys. He was a co-founder of a couple of startups, licensed some of his technologies to industry, and served as an Executive Advisor for Samsung Research.

Zheng Yan, photograph and biography not available at the time of publication.

Jia Liu (Senior Member, IEEE) is currently an Assistant Professor with the Center for Cybersecurity Research and Development, National Institute of Informatics, Tokyo, Japan. He has published more than 70 academic papers at premium international journals and conferences, such as IEEE TRANSACTIONS ON DEPENDABLEAND SECURE COMPUTING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, and IEEE INFOCOM. His research interests include wireless systems security, space-air-ground integrated networks, Internet of Things, and 6G. He received the IEEE Sapporo Section Encouragement Award, in 2016 and 2020.

Siwei Le, photograph and biography not available at the time of publication.